PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Semen quality of 1,559 young men from four cities in Japan
AUTHORS	Iwamoto, Teruaki; Nozawa, Shiari; Naka-Mieno, Makiko; Yamakawa, Katsunori; Baba, Katsuyuki; Yoshiike, Miki; Namiki, Mikio; Koh, Eitetsu; Kanaya, Jiro; Okuyama, Akihiko; Matsumiya, Kiyomi; Tsujimura, Akira; Kanetake, Hiroshi; Eguchi, Jiro; Skakkebæk, Niels; Vierula, Matti; Toppari, Jorma; Jørgensen, Niels

VERSION 1 - REVIEW

REVIEWER	Taisen Iguchi Professor National Institute for Basic Biology Japan I have no competing ineterests.
REVIEW RETURNED	26-Oct-2012

THE STUDY	This study does not include patients, therefore, one of above question is not applicable, but I checked No.
GENERAL COMMENTS	This manuscript contains very important data which can be used as a gold standard of semen quality in Japanese young men.

REVIEWER	Krzysztof Kula M.D.,Ph.D.,professor and chairmen, Dept. Andrology and Reproductive Endocrinology, Medical University of Lodz, Poland. No competing interests.
REVIEW RETURNED	13-Nov-2012

- **The reviewer completed the checklist but made no further comments.**

REVIEWER	Prof Richard M Sharpe Programme Group Leader MRC Centre for Reproductive Health University of Edinburgh, UK No competing interests
REVIEW RETURNED	12-Jan-2013

GENERAL COMMENTS	Whilst studying sperm counts and measuring reproductive hormone levels in young men is anything but new, the novelty of the present study is its setting –Japan. At a time of concern about poor/deteriorating reproductive health of the next generation, this

	study brings comparative data from the Asian scene to add to what we already have from Europe and the USA. The fact that it has used a similar standardized approach as has been used in Europe, also gives a measure of confidence when it comes to comparing data between the two regions. It is therefore worth publishing as it adds a clearer 'East-West' dimension to male reproductive health status. Having said all of this, it is clearly only useful to have such data if it is good data and its 'provenance' is clear – in other words, how much trust can be put in this data as truly reflecting the general population of young men in Japan. The authors address this in passing, but there are further issues that need addressing in this respect – I do not think that any discussion of biased recruiting etc will negate publication, rather it will make the data more transparent and alert readers to how much weighting to attach to particular results. Bias in recruitment. Unlike the European young men studies of military conscripts, the present studies rely wholly on young men volunteering. It is perhaps inevitable that those who 'think they may have some sort of reproductive problem' are more likely to take part. Of course, the great thing about sperm count is that it is completely hidden from the men themselves other than their awareness about size of their testes – the latter is only likely to raise concern when testes are very small, so the high average testis size in the present cohorts is suggestive that this is not an important factor – how did it compare with the partners of pregnant women? The authors could use this as evidence that sperm count is unlikely to have been subjected to bias. Table 1. The incidence of cryptorchidism seems high – is this self-reported, and if so was it verified from medical records, or was it present at examination (seems unlikely)? The high rate could be an indicator of recruitment bias. Similarly with orchitis, although this is only ~1%. The differences in hormone levels between centres is intriguing, albeit that the differences (in medians) are small. Did the authors examine the LH/T ratio to test for quality of Leydig cell function, and did this vary between centres? Was the % of men with an abnormal hormone profile (eg supranormal FSH or LH) different between centres? Page 12, lines 23-26. The authors state that 'the most important finding was that semen quality of the young men was poorer than that of partners of pregnant women'. But surely this is entirely as expected – and as all published such comparisons have found? The young men group will include infertile/subfertile men whereas the 'pregnant' group clearly excludes (or lessens) the contribution of these, so a higher sperm count is as expected. Page 11 line 17, p=0.05 (for difference in sperm morphology between groups) should be considered as significantly different as this is the arbitrary level set for significance. Page 11, lines 45-49. The numbers for varicocele on L and R do not add up to those diagnosed with a varicocele; does this mean some had a varicocele on both sides? The % who smoke is remarkably high by European standards – is this normal for Japanese students?

VERSION 1 – AUTHOR RESPONSE

Comment 1: Bias in recruitment. Unlike the European young men studies of military conscripts, the present studies rely wholly on young men volunteering. It is perhaps inevitable that those who 'think they may have some sort of reproductive problem' are more likely to take part. Of course, the great thing about sperm count is that it is completely hidden from the men themselves other than their awareness about size of their testes – the latter is only likely to raise concern when testes are very small, so the high average testis size in the present cohorts is suggestive that this is not an important factor – how did it compare with the partners of pregnant women? The authors could use this as evidence that sperm count is unlikely to have been subjected to bias.

Response: According to the above comments, we have added the following text to the Discussion.

'Since the sperm count cannot be known without laboratory analysis, it is unlikely that any of the volunteers would have had such information. Testicular size might hint to fertility problem if it were very small. However, there was only small difference (slightly larger in the present cohort) between the testicular size of the young and the fertile men15 in Japan. Moreover, semen variables of the healthy subgroup of young men (n=1,307) who had no history of reproductive problems were very similar to those of the entire study population (n=1,559).' (page 13, lines 15-25)

'Nevertheless, semen variables of the healthy subgroup of young men (Supplemental Table 1) who had no history of reproductive problems were very similar to those of the entire study population (Table 2). We thus considered that semen results of the present study population was unlikely to have been subjected to a significant bias.' (page 13, lines 52-54 & page 14, lines 9-11)

Comment 2: Table 1. The incidence of cryptorchidism seems high – is this self-reported, and if so was it verified from medical records, or was it present at examination (seems unlikely)? The high rate could be an indicator of recruitment bias. Similarly with orchitis, although this is only ~1%.

Response: The information about cryptorchidism was based on the questionnaire data, i.e. history of cryptorchidism, not current cryptorchidism. This included cryptorchidism at any time, i.e. both congenital and acquired, which may have different causes/consequences. Accordingly, we analysed the subgroup (n=1307) of the healthy men who had no history of cryptorchidism or any other reproductive problems, and compared that with the entire study population (n=1559). There were no differences in any of the semen variables, which is now mentioned in the Results (page 10, lines 45-49) as follows,

'These results were very similar to those from the healthy subgroup of young men (n=1,307: Supplemental Table 1) who had no history of reproductive problems, i.e. cryptorchidism, testicular torsion, orchitis, varicocele, inguinal hernia, STD, caused pregnancy, and experienced fertility problem.'

and also in the Methods, accordingly (page 6, lines 40-45).

'The healthy subgroup (n=1,307) of the entire study population (n=1,559), who had no history of cryptorchidism, testicular torsion, orchitis, varicocele, inguinal hernia, caused pregnancy, and experienced fertility problem, was examined separately (Supplemental Table 1).'

In addition, we have added the following text in the Discussion (page 13, lines 25-39).

‘Likewise, it should be noted that the incidence of cryptorchidism was considerably high, which could be an indicator of recruitment bias. In the present study, however, it is unlikely to be a recruitment bias, because the information about cryptorchidism was based on the questionnaire data, i.e. history of cryptorchidism, not current cryptorchidism. This included cryptorchidism at any time, i.e. both congenital and acquired, which may have different causes and consequences. In fact, there were no differences in semen quality between the healthy subgroup of young men (Supplemental Table 1) and the entire study population (Table 2) including the men with history of cryptorchidism or other reproductive problems.’

Comment 3: The differences in hormone levels between centres is intriguing, albeit that the differences (in medians) are small. Did the authors examine the LH/T ratio to test for quality of Leydig cell function, and did this vary between centres? Was the % of men with an abnormal hormone profile (eg supranorma FSH or LH) different between centres?

Response: The number of men with high (mean+2SD) gonadotropin levels was very small (1-4 men for FSH 0.3-1.3%, and 2-14 for LH 0.7-4.7%), and that gave poor statistical power to evaluate any centre differences for these measurements. Thus there were no statistically significant differences in the number of men with high FSH, whereas there was a difference in the number of men with high LH. However, we doubt whether this is biologically significant. T/LH ratio varied between the centres except for Osaka vs. Nagasaki. In Kawasaki, the ratio was highest (9.74±4.26), while in Osaka, Nagasaki, and Kanazawa the values were 8.75±3.54, 8.36±3.76, and 7.45±3.51, respectively. All pair-wise comparisons except Osaka vs. Nagasaki were statistically significant. Accordingly, we have added the following text to the Results,

‘The number of men with high (mean+2SD) gonadotropin levels was very small (1-4 men for FSH 0.3-1.3%, and 2-14 for LH 0.7-4.7%). T/LH ratio of men from Kawasaki, Osaka, Kanazawa, and Nagasaki were 9.74±4.26, 8.75±3.54, 8.36±3.76, and 7.45±3.51, respectively. All pair-wise comparisons between centres except Osaka vs. Nagasaki were statistically significant ($p<0.0001$-0.01).’ (page 11, lines 35-44)

and to the Discussion.

‘T/LH ratio may inform about Leydig cell function, and therefore we also analyzed it. Variation between the study centres was small but statistically significant; however, biological significance of variation in this scale is questionable.’ (page 14, lines 27-30)

Comment 4: Page 12, lines 23-26. The authors state that ‘the most important finding was that semen quality of the young men was poorer than that of partners of pregnant women’. But surely this is entirely as expected – and as all published such comparisons have found? The young men group will include infertile/subfertile men whereas the ‘pregnant’ group clearly excludes (or lessens) the contribution of these, so a higher sperm count is as expected.

Response: We have modified the text accordingly: the following text in the Discussion ‘The most important finding is that’ has been changed to ‘As expected’. (page 12, line 35)

Comment 5: Page 11 line 17, $p=0.05$ (for difference in sperm morphology between groups) should be considered as significantly different as this is the arbitrary level set for significance.

Response: We agree with this comment and revised the text accordingly.

The text ‘total number of morphologically normal spermatozoa (p<0.0001) than young men, while the percentage of normal spermatozoa did not differ between the groups (p=0.05)’ has been changed to:

‘total number of morphologically normal spermatozoa (p<0.0001), and the percentage of normal spermatozoa (p=0.05) than young men’ (page 11, lines 17-20)

Comment 6: Page 11, lines 45-49. The numbers for varicocele on L and R do not add up to those diagnosed with a varicocele; does this mean some had a varicocele on both sides?

Response: There were bilateral cases and we simplified the text in the Results (page 12, lines 10-15) to make calculations clearer as follows.

‘Varicocele was on the left side only in 23.6%, both sides in 2.9% and only on the right side in 0.6% of men’

Comment 7: The % who smoke is remarkably high by European standards – is this normal for Japanese students?

Response: The smoking rate in Japan has been decreasing from the 1980s, but it is still very high for a developed nation. The National Health and Nutrition Survey (Ministry of Health, Labour and Welfare, Japan) showed that the smoking rates of Japanese men in 2000 were 47.4% in total and 60.8% in 20-29 age group. These figures seem to be consistent with the results from the Japanese students. Therefore, we have added the following text to the Discussion (page 156, lines 35-42).

‘By contrast, the smoking rates of the young men themselves looks exceedingly high for a developed nation, but this is a common trend in Japanese male, which is consistent with the figures from National Health and Nutrition Survey (47.4% in total and 60.8% in 20-29 age group in 2000)28, as well as the results from the previous our study of the fertile Japanese men (52.8% in total) 15’

Other corrections

Relating to the above revisions, the following two new references have been added to the References and have cited in the corresponded parts as reference numbers 15 and 28.

15. Teruaki Iwamoto, Shiari Nozawa, Miki Yoshiike, et al. Semen quality of fertile Japanese men: a cross-sectional population-based study of 792 men. BMJ Open 2013;3: e002223.

28. The National Health and Nutrition Survey 2008, Ministry of Health, Labour and Welfare, Japan

Minor correction: We have changed ‘April’ to ‘May’ (page 6, line 30; simple mistake)

VERSION 2 – REVIEW

REVIEWER	R Sharpe University of Edinburgh
REVIEW RETURNED	18-Mar-2013

GENERAL COMMENTS	The authors have satisfactorily addressed all of my comments