Degenerative cervical spine changes among early career fighter pilots: a 5-year follow-up
Tuomas Keskimölö,1 J Pernu,1 J Karppinen,2,3 J Niinimäki,1,2 P Oura,2 T Leino,4,5 T Honkanen6

ABSTRACT

Introduction Degenerative changes of the cervical spine often cause disability and flight duty limitations among Finnish Air Force (FINAF) fighter pilots. We aimed to study the effect of +Gz exposure on degenerative changes in the cervical spine by comparing cervical MRIs of FINAF fighter pilots and controls.

Methods At baseline, the volunteer study population consisted of 56 20-year-old FINAF male fighter pilots (exposure group) and 56 21-year-old Army and Navy cadets (control group). Both groups underwent MRI of the cervical spine at the baseline and after 5 years. Degenerative changes evaluated using MRI included intervertebral disc (IVD) degeneration (Pfirrmann classification), disc herniations, uncovertbral arthrosis, Schmor’s nodes, Modic changes, spinal canal stenosis, kyphosis and scoliosis.

Results The degree of IVD degeneration in the whole cervical spine increased significantly in both populations with no between-group differences. The prevalence of disc herniations also tended to increase in both populations with no difference in the incidence over the follow-up. However, pilots proved to have more disc herniations at the baseline and at the follow-up. There were virtually no between-group differences in other assessed degenerative changes.

Discussion We found that IVD degeneration and the prevalence of disc herniations increased at a similar rate for fighter pilots and non-flying military students when all cervical levels were summed up. The lack of difference may be explained by the relatively low cumulative +Gz exposure during the first 5 years of a pilot’s career.

INTRODUCTION

Fighter pilots need to move their heads with full range of motion in three directions to observe the airspace around them with simultaneous exposure to high +Gz forces.1 The combination of +Gz forces and awkward head positions is potentially harmful for the structures of the cervical spine.2 Neck pain is reported to be a common health problem among military pilots.3,4 A meta-analysis, consisting of 18 studies and more than 8000 pilots, revealed that 51% of fighter pilots suffered from neck disorders, and one-third of them had sought medical care for neck pain.5 Degenerative changes of the cervical spine in MRI include decreased signal intensity and height of intervertebral disc (IVD), annular fissures, protrusions/herniations and bone marrow lesions.6 In a meta-analysis by Shiri et al.,6 neck pain was more common among pilots exposed to high +Gz-loads compared with pilots flying less +Gz-capable aircraft, but the authors found no difference between the groups in terms of cervical IVD degeneration. However, the results of individual studies on cervical IVD degeneration are conflicting. For example, Hämäläinen and Petren-Mallmin and Linder found that there was an increased prevalence of degenerative changes in MRI among fighter pilots when compared with age-matched controls.7,8 Petren-Mallmin and Linder8 found that experienced military pilots had significantly more osteophytes, protrusions/herniations and both spinal and foraminal stenosis. Furthermore, Hendriksen and Holewijn6 compared F-16 pilots to fighter pilots flying less G-capable aircraft and found that frequent exposure to high +Gz forces may cause premature degenerative changes. However, several other studies10–12 did not observe a significant difference in cervical IVD degeneration between the pilots and controls.

The contributing factors to fighter pilots’ IVD degeneration are suggested to be the head movements during the Gz exposure and cumulative trauma thereof. A recent study investigated in-flight neck muscle activity during combat manoeuvring. The results implicated that high +Gz combined with awkward head postures (eg, rotations in combination with extensions) causes a force exceeding the load-bearing capacity of the cervical muscles.1 This possibly results in more loading to the bony structures, IVDs and ligaments of the spine and may be a risk factor for early IVD degeneration. Furthermore, it has been suggested that the cumulative soft tissue (muscle or ligament) injuries may lead to
unstable spinal structures that lead to abnormal IVD loads and later to degeneration.

Neck disorders may lead to temporary or permanent flight disqualifications or limitations. In the Finnish Air Forces (FINAF), disability accompanied by spinal IVD degeneration is reported as the most common reason for aeromedical limitations. These flight duty limitations do not only cause individual harm but are economically and operationally significant. Fighter pilot training is highly expensive, which limits the number of pilots selected for fighter training. In order to preserve a sufficient number of fighter pilots, the pilots must maintain their working capability both physically and mentally. Therefore, it is important to take care of and monitor the health of fighter pilots already in the training phases.

The primary aim of the present study was to investigate the effect of +Gz exposure on degenerative changes of the cervical spine by comparing FINAF fighter pilots and controls. The degenerative changes of the cervical spine (from C2 to Th1) were assessed by MRI.

METHODS

Subjects

At baseline, the volunteer study population consisted of 56 FINAF male fighter pilot cadets and 56 sex-matched Finnish Defense Force Army and Navy cadets from the National Defense University as controls. FINAF cadets were flying BAe Hawk jet trainer (up to +8 Gz) during the follow-up. Whereas the control group’s occupations during follow-up varied mainly from office work to conscript trainers that do not expose to accessory neck strain or G force stress. All pilots and controls completed follow-up. A questionnaire considering health and physical condition was filled by both groups. The questionnaire included questions on age, weight, height and physical performance. This structured questionnaire has been used previously. We used the STROBE cohereporting guidelines for this study (online supplemental file 1).

The results of the first Cooper’s test (12 min maximal distance running test) as a conscript were used to determine the physical condition of the subjects. Before starting the research, the procedures were explained to all the participants, who gave their voluntary written informed consent under the Declaration of Helsinki.

MRI examination

The MRI examinations were performed at the baseline and after a 5-year follow-up for both populations, between 2008 and 2020. MRI images were acquired using clinical 1.5T MRI scanners. The imaging protocols included conventional T1-weighted and T2-weighted fast spin echo sequences: for example, 11 ms TE (Echo Time) and 440 ms TR (Repetition Time) for T1-weighted and 113 ms TE and 3300 ms TR for T2-weighted sagittal images and 100 ms TE and 3700 ms TR for T2-weighted axial images. The slice thickness was 3 mm, and the in-plane resolution was 0.47×0.47 mm for the sagittal plane and 0.35×0.35 mm for the axial plane.

The degree of IVD degeneration was assessed using the Pfirrmann classification with five grades consisting of evaluations of the disc space height, disc homogeneity, nucleus intensity and distinction between nucleus and annulus. An IVD degeneration sum score was composed by adding the scores of degenerated discs (Pfirrmann grades 1 and 2=0 p, grade 3 (mild degeneration)=1 p, grade 4 (moderate degeneration)=2 p, grade 5 (severe degeneration)=3 p) together and divided by the number of subjects. IVD herniations were classified as protrusions, extrusions and sequesters but were later combined due to the low number of extrusions and sequesters. Schmorl’s nodes were evaluated from each vertebral body throughout the cervical spine. Osteophytes were assessed from both sides of vertebral bodies and classified as uncovertebral arthrosis whether the osteophytes were unilateral or bilateral. Scoliosis was noted if a scoliotic angle of at least 10° was measured. The spinal rotation was not measured as we did not address a medical diagnosis of scoliosis. The kyphotic angle was measured using the Harrison posterior tangent method and recorded if the kyphotic angle in the cervical spine was at least 1°.

RESULTS

Subjects

There was no significant difference at the baseline in height, weight, BMI or last school grade of physical education between the groups. However, the pilots were younger and had performed better in their first Coopers test as a conscript (Table 1).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Pilots (n=56)</th>
<th>Controls (n=56)</th>
<th>P for difference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>20 (±0.70)</td>
<td>21 (±1.2)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>179.9 (±4.7)</td>
<td>180.4 (±5.6)</td>
<td>0.650</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>75.7 (±6.3)</td>
<td>77.0 (±7.7)</td>
<td>0.346</td>
</tr>
<tr>
<td>BMI (kg/m²)</td>
<td>23.4 (±1.6)</td>
<td>23.6 (±1.9)</td>
<td>0.467</td>
</tr>
<tr>
<td>Cooper’s test* (m)</td>
<td>3040 (±190)</td>
<td>2780 (±420)</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Last physical education grade</td>
<td>9.29 (±0.62)</td>
<td>9.13 (±0.85)</td>
<td>0.258</td>
</tr>
</tbody>
</table>

The values are means with SD. BMI, body mass index.
Modic change during the follow-up. There was no statistically significant difference in height, weight, BMI or physical education grade between the groups. The difference in age proved to be statistically significant, but the difference was only 0.7 years. The notable difference in the average result of the Cooper’s test may be partly explained by more demanding physical performance requirements during the application phase for the fighter pilot training.

Comparison to previous literature
The present findings of IVD degeneration were similar to the meta-analysis by Shiri et al.6 However, when comparing our results with individual studies, there were some conflicting results,7,8 whereas others were well in line with our results.9,10 When comparing the present results to previous studies, it has to be kept in mind that the previous investigations have had limitations such as a small number of subjects. Both Hämäläinen and Sovelius used only 12 fighter pilots with 12 age-matched controls.7,11 In Sovelius’s study, degenerative changes in the cervical spine of fighter pilots were more common in the lower part of the cervical spine, especially at the C5/C6 and C6/C7 levels.11 Instead, the degenerative changes in the control group were more scattered in both studies.7,11

The fighter pilots had a total experience of 5 years of military flying and only 3 years of flying with high performance aircraft (up to +8 Gz). A comparison between our study and other studies must be made with caution, since the Hawk flight hours were below 200.

Magnetic resonance imaging
MRI is widely used in this research field. The Harrison posterior tangent method is an accurate method for estimating the curvature of the cervical spine, and it has lower SEs of measurement than the four-line or two-line Cobb methods.22 However, our intention was not to diagnose kyphosis but only to assess kyphotic angles in the images. MRI also shows the shape of IVDs and vertebrae clearly and is commonly used for recording Schmorl’s nodes.19

The prevalence of some MRI findings (eg, Modic changes and spinal canal stenosis) was low, excluding any prominent differences between the groups. The inter-rater disagreement in IVD degeneration is endorsed by the fact that the Pfirrmann classification is primarily structured for and used in the assessment of lumbar IVDs, but it has been adapted also in assessment of cervical spine.16,25 Our aim was to assess the incidences between groups and not primarily the prevalences.

When comparing results with previous studies, it must be taken into account that they have used rather low-quality MRI devices. Hämäläinen used a 0.1 T MRI device, and Petren-Mallmin and Linder used a 0.5 T device throughout the study.7,8 Sovelius used a 0.1 T MRI device in the baseline phase, which was replaced by a 1.0 T device in follow-up phase.11 In the Hendriksen and Holewijn’s study, the mean age of the F-16 group was higher than the control group. After the groups were adjusted for age, the difference between the groups disappeared.

When comparing IVD degeneration of our population to civilian 21–30-year-old male population, the results were similar with Pfirrmann grades 2 and 3 being clearly the most common.25 Our finding of the highest prevalence of disc herniations and IVD degeneration in the lower cervical spine (ie, C5/C6) is also similar to other studies. The lower cervical spine was most affected in non-flying 12–95-year-old population and age-matched FINAF fighter pilots.11,25

IVD degeneration
During the follow-up, a statistically significant progression in IVD degeneration was found in both groups. Progression was noted for both pilots and controls in IVD degeneration at all levels from C2/C3 to C6/C7 (Table 2) and in the IVD degeneration sum score (Table 3). At C7/Th1, the change was not significant in either group. When comparing the groups, the change in Pfirrmann score among the fighter pilots was of greater magnitude at C6/C7 (p=0.007). There was no significant difference in the change of the IVD degeneration sum score between the groups. There was no between-group difference at the baseline or follow-up (Table 3, Figure 1).

IVD herniations
The total number of IVD herniations increased significantly among controls (p=0.004) and showed an almost increasing trend among pilots (p=0.061). Both at baseline and follow-up, the fighter pilots had a higher prevalence of IVD herniations, but there was no difference in the incidence of IVD herniations between the groups (Table 3).

Other cervical degenerative changes
The prevalence of uncovertebral arthrosis (number of arthritic vertebral levels per individual) did not differ at baseline, but at follow-up, it was higher among the pilots. There was no difference in the incidence of uncovertebral arthrosis between the groups. No between-group differences were observed in Schmorl’s nodes, Modic changes, kyphosis, spondylolisthesis or stenosis at the baseline or follow-up. Two pilots developed type 1 Modic change during the follow-up, whereas one control had a type 2 Modic change at the baseline and developed one more type 2 Modic change during the follow-up (Table 3).

Rater reliability
All intrarater reliabilities were substantial or perfect (0.6–1.0). Inter-rater reliabilities were between 0.6 and 1.0 for disc herniations, Schmorl’s nodes, Modic changes, kyphosis and spondylolisthesis. There was more disagreement in uncovertebral arthrosis (0.3) and IVD degeneration (0.2) in which the disagreement was linear in nature.

DISCUSSION
According to the present findings, there was a progression of IVD degeneration in both groups. Additionally, the prevalence of disc herniations increased significantly among the controls and had an increasing trend of borderline significance among pilots. However, a more prominent progression in the fighter pilots with cumulative +Gz exposure was observed only in IVD degeneration at C6/C7. We did not find any progression of uncovertebral arthrosis, Schmorl’s nodes, Modic changes, kyphosis, spondylolisthesis or spondylolisthesis during the follow-up. We found higher prevalence of herniations in pilots at baseline and at follow-up. Because either of the populations had been exposed to G force stress at baseline and the other underlying factors are unknown, our primary aim was to assess the incidences. Also, the controls seem to trend worse in disc degeneration sum score medians at baseline (Table 3). However, the difference was not significant either in disc degeneration sum score medians or in average Pfirrmann scores (pilots 2.22, controls 2.33).

Control population selection
The control population proved to be physically very similar to the study population. There was no statistically significant difference in height, weight, BMI or physical education grade between the groups. The difference in age proved to be statistically significant, but the difference was only 0.7 years. The notable difference in the average result of the Cooper’s test may be partly explained by more demanding physical performance requirements during the application phase for the fighter pilot training.
Table 2  Prevalence of MRI findings among pilots and controls

<table>
<thead>
<tr>
<th>Finding</th>
<th>Baseline</th>
<th>Follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>C2/3</td>
<td>C3/4</td>
</tr>
<tr>
<td>Pilots</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pfirrmann grade ≤2</td>
<td>66.1 (37)</td>
<td>58.9 (33)</td>
</tr>
<tr>
<td></td>
<td>33.9 (19)</td>
<td>33.9 (19)</td>
</tr>
<tr>
<td>Pfirrmann grade ≥4</td>
<td>0.0 (0)</td>
<td>7.1 (4)</td>
</tr>
<tr>
<td>Disc herniation</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Uncovertebral arthrosis</td>
<td>0.0 (0)</td>
<td>3.6 (2)</td>
</tr>
<tr>
<td>Schmorl's node</td>
<td>1.8 (1)</td>
<td>1.8 (1)</td>
</tr>
<tr>
<td>Modic change</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Spinal canal stenosis</td>
<td>0.0 (0)</td>
<td>0.0 (0)</td>
</tr>
<tr>
<td>Kyphosis</td>
<td>30.4 (17)</td>
<td></td>
</tr>
<tr>
<td>Scoliosis</td>
<td>3.6 (2)</td>
<td></td>
</tr>
</tbody>
</table>

The values are percentages (number of findings per individual) with frequencies (number of findings at each cervical level).
The association between spinal injury-induced flight duty limitations and cumulative +Gz exposure in FINAF pilots was studied in 2018. The cumulative +Gz stress of the first 5 years of a pilot's career was assessed using a fatigue index. The study population consisted of 23 pilots with spinal injury-induced flight duty limitations and the controls of 50 symptomless pilots. At the time (between 1995 and 2004), the cumulative +Gz load was much higher due to the old fighter training programmes. Thus, our results may be partially explained by the relatively low +Gz exposure during the first 5 years of the pilots' careers. Furthermore, the 5-year follow-up period of the young fighter pilots may be too short to discern a difference between flying and non-flying personnel.

Study evaluation
The strength of present study was the prospective study design and the physically similar groups. In contrast to previous studies, the considerable number of fighter pilots included in this study increased the reliability of our findings. Furthermore, the use of reliable and valid imaging modality (MRI) in this investigation enhances the quality of the study. We classified the MRI findings according to validated methods, which enhances the comparability of our results to other studies.

CONCLUSIONS
In summary, this was a novel prospective study with proper control selection to evaluate the effects of +Gz forces on degenerative changes in the cervical spines of fighter pilots. The result of our study indicates that the effects of +Gz forces on cervical spine degenerative changes were not statistically significant for the most part in comparison with non-flying subjects. The only significant difference between the groups was in the incidence of IVD degeneration of the pilots at C6/C7. It is noteworthy that our fighter pilots were exposed to high +Gz forces for only 3 years during the follow-up period.

The present 5-year follow-up study suggests that significant cervical degeneration is prevalent in young healthy male military personnel. Therefore, further research is needed to ensure that unnecessary neck strain is not applied to fighter pilots or other military personnel.

Acknowledgements We would like to thank Hannu Vähänikkilä, PhD, for advice in statistical analysis.

Contributors The manuscript has been read and approved by all authors. TK has broadly taken part in statistical analysis and drafting and composing the article and also submitted the study. JP has taken big part in drafting and composing the article, especially in the introduction and discussion sections considering previous literature and analysing the data. JK has largely revised and edited the article and taken big part in the interpretation of results. JN took part in defining the study design and analysing the data. JK has largely revised and edited the article and taken big part in drafting and composing the article and analysing the data. TL greatly revised and edited the article and took big part in statistical analysis. JP has taken big part in drafting and composing the article.

Funding This study was financially supported by a research grant from the Centre for Military Medicine, Finnish Defence Forces, Finland.

Competing interests None declared.

Patient consent for publication Not required.

Ethics approval The study protocol was approved by the Ethical Committee of the Central Finland Health Care District.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement No data are available. All data according to Finnish Air Force fighter pilots is considered as classified, and therefore, data used in this study are not available.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) licence, 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 and the nature of the licence is indicated. You may access the full text for this article here: http://militaryhealth.bmj.com/
REFERENCES


Reporting checklist for cohort study.

Based on the STROBE cohort guidelines.

**Instructions to authors**

Complete this checklist by entering the page numbers from your manuscript where readers will find each of the items listed below.

Your article may not currently address all the items on the checklist. Please modify your text to include the missing information. If you are certain that an item does not apply, please write "n/a" and provide a short explanation. Please find explanations on the last pages of the paper.

Upload your completed checklist as an extra file when you submit to a journal.

In your methods section, say that you used the STROBE cohortreporting guidelines, and cite them as:


<table>
<thead>
<tr>
<th>Reporting Item</th>
<th>Page Number</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Title and abstract</strong></td>
<td></td>
</tr>
<tr>
<td>Title</td>
<td>#1a</td>
</tr>
<tr>
<td>Abstract</td>
<td>#1b</td>
</tr>
<tr>
<td><strong>Introduction</strong></td>
<td></td>
</tr>
<tr>
<td>Background / rationale</td>
<td>#2</td>
</tr>
<tr>
<td>Objectives</td>
<td>#3</td>
</tr>
<tr>
<td><strong>Methods</strong></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td>#4</td>
</tr>
<tr>
<td>Setting</td>
<td>#5</td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>----</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>#6a</td>
</tr>
<tr>
<td>Eligibility criteria</td>
<td>#6b</td>
</tr>
<tr>
<td>Variables</td>
<td>#7</td>
</tr>
<tr>
<td>Data sources / measurement</td>
<td>#8</td>
</tr>
<tr>
<td>Bias</td>
<td>#9</td>
</tr>
<tr>
<td>Study size</td>
<td>#10</td>
</tr>
<tr>
<td>Quantitative variables</td>
<td>#11</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>#12a</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>#12b</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>#12c</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>#12d</td>
</tr>
<tr>
<td>Statistical methods</td>
<td>#12e</td>
</tr>
</tbody>
</table>

**Results**

Participants #13a Report numbers of individuals at each stage of study—eg numbers potentially eligible, examined for eligibility, confirmed eligible, included in the study, completing follow-up, and analysed. Give information separately for exposed and unexposed groups if applicable.

Participants #13b Give reasons for non-participation at each stage n/a

Participants #13c Consider use of a flow diagram n/a

Descriptive data #14a Give characteristics of study participants (eg demographic, clinical, social) and information on exposures and potential confounders. Give information separately for exposed and unexposed groups if applicable.

Descriptive data #14b Indicate number of participants with missing data for each variable of interest n/a

Descriptive data #14c Summarise follow-up time (eg, average and total amount), page 4

Outcome data #15 Report numbers of outcome events or summary measures over time. Give information separately for exposed and unexposed groups if applicable. page 6

Main results #16a Give unadjusted estimates and, if applicable, confounder-adjusted estimates and their precision (eg, 95% confidence interval). Make clear which confounders were adjusted for and why they were included

Main results #16b Report category boundaries when continuous variables were categorized n/a

Main results #16c If relevant, consider translating estimates of relative risk into absolute risk for a meaningful time period, n/a
### Other analyses  
**#17** Report other analyses done—eg analyses of subgroups and interactions, and sensitivity analyses  

### Discussion

**Key results  
**#18** Summarise key results with reference to study objectives  

**Limitations  
**#19** Discuss limitations of the study, taking into account sources of potential bias or imprecision. Discuss both direction and magnitude of any potential bias.  

**Interpretation  
**#20** Give a cautious overall interpretation considering objectives, limitations, multiplicity of analyses, results from similar studies, and other relevant evidence.  

**Generalisability  
**#21** Discuss the generalisability (external validity) of the study results  

### Other Information

**Funding  
**#22** Give the source of funding and the role of the funders for the present study and, if applicable, for the original study on which the present article is based  

None

The STROBE checklist is distributed under the terms of the Creative Commons Attribution License CC-BY. This checklist can be completed online using [https://www.goodreports.org/](https://www.goodreports.org/), a tool made by the [EQUATOR Network](https://www.equator-network.org/) in collaboration with [Penelope.ai](https://www.penelope.ai).

6a, 10: All fighter pilot cadets from particular years were included and the number of controls was matched.

12b: There were no meaningful subgroups to address.

12c,d: There was no missing data

12e: None was conducted

13b, 14b: All pilots and controls completed follow-up

16a: No meaningful assessable confounders were available

16b: Continuous categorical variables were not categorized

16c: Not relevant
17: No other analyses

22: The study has not acquired external funding