Plastic and Reconstructive Surgery Advance Online Article

DOI: 10.1097/01.prs.0000438455.37604.0f

A systematic review and meta-analysis on the prevalence of Dupuytren Disease in the general population of western countries

Rosanne Lanting[1], MD, Dieuwke C. Broekstra[1], MSc, Paul. M.N. Werker[1], MD, PhD, Edwin R. van den Heuvel[2], PhD

[1] University of Groningen, University Medical Center Groningen, Department of Plastic Surgery, Groningen, The Netherlands
[2] University of Groningen, University Medical Center Groningen, Department of Epidemiology, Groningen, The Netherlands

R. Lanting and D.C. Broekstra contributed equally to this article.

Addresses:
[1] University Medical Center Groningen
Department of Plastic Surgery BB81
P.O. Box 30 001
9700 RB Groningen
The Netherlands

[2] University Medical Center Groningen
Department of Epidemiology FA40
P.O. Box 30 001
9700 RB Groningen
The Netherlands

Meetings at which the paper has been presented:
None.

Running head: Prevalence of Dupuytren Disease

Corresponding author:
R. Lanting, MD
Department of Plastic Surgery
HPC BB81
P.O. Box 30.001
9700 RB Groningen
The Netherlands
Tel.: +31-50-3618840
Fax: -
E-mail: r.lanting@umcg.nl

Financial disclosure
Author P.M.N. Werker is a consultant for the pharmaceutical company Pfizer. Author E.R. van den Heuvel is a consultant for Merck (in our opinion this does not cause a conflict of interest).
The other authors have no financial or other interest in any of the products, devices, or drugs mentioned in this manuscript.

Funding:
This research was funded by the University Medical Center Groningen.

Author's role:

R. Lanting:
- Conception and design of the study.
- Data collection and analysis.
- Writing article and revising article.
- Approval of the final version of the article.

D.C. Broekstra:
- Conception and design of the study.
- Data collection and analysis.
- Writing article and revising article.
- Approval of the final version of the article.

P.M.N. Werker:
- Conception and design of the study.
- Revising article.
- Approval of the final version of the article.

E.R. van den Heuvel:
- Conception and design of the study.
- Data analysis.
- Writing article and revising article.
- Approval of the final version of the article.

Abstract

Background

Dupuytren Disease (DD) is a fibroproliferative disease of palmar fascias of the hand. Prevalence of DD has been the subject of several reviews, though an accurate description of the prevalence range in the general population -and of the relation between age and DD- is lacking.

Methods

A systematic review was performed by searching Embase and Pubmed on database specific mesh terms, and in title and abstract for “Dupuytren”, “incidence”, and “prevalence”. Two reviewers independently assessed the papers using inclusion and exclusion criteria, and rated the included studies with a quality assessment instrument. In a meta-analysis the median prevalence, as function of age by gender, was estimated, accompanied with 95% prediction intervals. The observed heterogeneity in prevalence was investigated with respect to quality of the studies and geographical location.

Results

Twenty-three of 199 unique identified papers were included. Number of participants ranged from 37 to 97,537, aging 18-100 years. Prevalence varied from 0.6-31.6%. The quality of studies differed, but could not explain the heterogeneity between studies. Mean prevalence was estimated 12%, 21%, and 29% at ages 55, 65, and 75 respectively, based on the relationship between age and prevalence determined from ten studies.

Conclusions

We were able to describe a prevalence range of DD in the general population of western countries. Furthermore, the relationship between age and prevalence of DD is given per gender, including 95% prediction intervals. Hereby, it is possible to determine the prevalence at a certain age for the total population, and for men and women separately.

Level of Evidence: Prognostic/Risk Studies, III

Introduction

Dupuytren Disease (DD) is a fibroproliferative disease which affects the palmar fascias of the hand. This results in the development of nodules and cords, which eventually may contract and give rise to flexion contractures of the affected fingers.

The origin of DD has been attributed to both genetic and environmental factors. The results of several family studies, and more specific twin studies, suggested that DD has a strong genetic component.[1-3] In 2011, Dolmans et al. performed a genome wide association study in which nine genes that are associated with DD were identified.[4]

Some environmental risk factors include excessive alcohol consumption, smoking, manual work and hand trauma.[5, 6] In addition, several diseases, such as diabetes mellitus and epilepsy, are thought to play a role in the etiology of DD.[7-9] However, the role of these risk factors and diseases is not fully cleared, and the results of different studies are occasionally conflicting.

Many articles about the prevalence of DD have been published.[10-15] In these articles there is a wide range of prevalence rates, varying from 0.2% to 56%[16, 17], as reported in a previous literature review.[18] This wide range, in our opinion, may at least partly be caused by the great heterogeneity between study populations, such as healthy populations, participants with certain risk factors as well as patients with specific diseases. Suboptimal design of the included studies may also be a reason for the wide range.

Until now, no systematic review was conducted to scrutinize the prevalence rates specifically in the general population, i.e. a healthy non-hospital population. It is assumed that life expectancy will increase considerably in the coming decades[19], and from our clinical experience we know that DD is a chronic disease of the elderly. Therefore, it will be important to enhance our knowledge about prevalence rates in the general population, and to be aware of changes in the prevalence across age. Furthermore, new treatment options have emerged, such as radiotherapy, percutaneous needle fasciotomy, and collagenase injection, and prevalence rates may be used to evaluate their cost effectiveness.

The aim of this study was to specify the prevalence range of DD in the general population, i.e. a healthy non-hospital population. This was done by reviewing the literature on prevalence of DD

systematically, combined with a quality assessment of the included studies. A secondary goal was to perform a meta-analysis on the relation between age and prevalence of DD.

Methods

Literature search

Our final literature search was performed on 9th of May, 2012 in the bibliographical databases PubMed and Embase, since earlier searches in the databases “Web of Science” and “Cochrane Library” had not retrieved any additional results. PubMed was searched with the search strategy: ("Dupuytren Contracture"[Mesh] OR dupuytren*[TIAB]) AND ("Prevalence"[Mesh] OR prevalen*[TIAB] OR "Incidence"[Mesh] OR "incidence"[TIAB]). In Embase the following search strategy was imputed: dupuytren*:ab,ti AND ('prevalence'/exp OR prevalen*:ab,ti OR 'incidence'/exp OR 'incidence':ab,ti) NOT [medline]/lim AND [embase]/lim.

The search was updated on 24th of January, 2013, and was supplemented by automatically weekly derived updates from PubMed until 4th of August. No limits were implemented in our search queries.

Assessment of relevant studies

Two authors (RL and DB) independently assessed the studies in three rounds, based on predefined criteria (Table 1), and Cohen’s kappa was calculated for each round. If in the first round inclusion or exclusion criteria could not be assessed from the title and abstract, a full text analysis was performed. After each round, discrepancies were discussed to reach consensus. The third author (PW) was consulted if no consensus could be reached.

Quality assessment of included studies

We used the scoring instrument of Cho[20] to assess the quality of the studies, based on a review article on quality assessment tools for epidemiologic studies.[21]

The instrument consists of 24 questions about study design, participants, methods to control bias, statistical analyses, reporting of results, and the conclusions drawn from the results.

For each question respectively 2, 1, 0, and 0 points were awarded to the answers "Yes," "Partial," "No," and "Not applicable", in order to obtain an overall quality score for each article.

This was done for each question except for the question on study design; in that case 1 to 5 points were given (1 for case reports, 2 for time series or uncontrolled experiments, 3 for cohort or case-control studies, 4 for nonrandomized control trials, and 5 for randomized control trials).[20]

Total points awarded for the 24 questions were divided by the total possible points (the sum of the maximum points for each item, excluding "Not applicable" items) to generate a fraction between 0 and 1. A score of 1 represents the highest quality.[20]

All articles that were included after the second full text round were scored with this instrument by RL and DB independently. The article by Lanting et al.[22] was evaluated by DB and an independent clinical epidemiologist to avoid a conflict of interest.

Data extraction and statistical analysis

In a statistical analysis, we combined studies that provided information on prevalence and sample sizes for different age categories in a total population, or in males and females separately. The aim of this meta-analysis was to determine a population-averaged relationship between age and DD, and to study possible heterogeneity in this relationship between studies. The mid points of the age categories were used in a generalized linear mixed model. The form of the age-prevalence relationship was selected equal to an asymmetric logistic function with a random intercept for study to address possible heterogeneity. This model was applied to the data of males and females simultaneously with a random intercept for males and females that was correlated. A simpler model with only one random intercept was applied to the totals of males and females, since some studies did not provide data separately by gender. From the estimated models and the random effects, a range of age based predicted prevalences were estimated (i.e. 95% prediction intervals). Additionally, in case heterogeneity was present, it was investigated whether the overall quality score, the quality of study design or geographical location affects the heterogeneity.

In some of the studies, the prevalence was determined in patients with a specific disease, and in a control group. If that was the case, only the data from the control group were used. The calculation of the exact 95% confidence intervals for the overall proportion of DD was calculated using the F-distribution.[23]

Results

Results of literature search and assessment of relevant studies

The search resulted in 212 articles. After excluding duplicates and critical appraisal of the studies by predefined criteria (Table 1), 23 studies were included (Figure 1).Two main reasons led to exclusion: first, the prevalence of DD was not determined, and second, the study population was not a general population. As a consequence also all non-English papers were excluded. To quantify the decisions in the selection process, we performed a Cohen's kappa analysis for each round of assessment; title and abstract (κ = 0.623, $p < 0.001$); full text round 1 ($\kappa = 0.449$, $p = 0.001$); and full text round 2 ($\kappa = 0.528$, $p = 0.001$).

As shown in Table 2, articles were published between 1972 and 2013. In some studies, only data from the control group were used (noted as CG in Table 2). Several times these control groups were chosen from a population that sustained hand pathology.[24-26] In two studies it was explicitly noted that the control group did not suffer from hand pathology.[27, 28]

The total number of participants in the included studies ranged from 37 to 97,537, and in seven of these studies only males participated.[29-35] Age ranged from 18 to 100 years, with an average above 50 years in 12 studies. In six studies age was only reported in categories, without absolute number of participants in each category, so it was not possible to calculate a mean age (CAT in Table 2).[26, 27, 34, 36-38]

The lowest prevalence found was 0.6% compared with 31.6% as highest prevalence over all age groups.[12, 39] In two studies, DD was diagnosed in a different fashion compared with the other studies. Descatha et al. did not include palmar thickening as sign of DD, and Lucas et al. excluded the thumb from examination.[32, 33]

The quality score is depicted in the last column of Table 2, this score ranged from 0.23 to 0.80.

Results of quality assessment

Table 3 shows in detail the results of the quality assessment per question, and Table 4 shows the score on the different questions per study. Question 2 is an open question which does not contribute to the final score.

The majority of studies reported the study question only partially. In 13% of the studies the inclusion and exclusion criteria were completely explained, while in 61% these criteria were not depicted at all. In almost 80% of the papers, the subjects were not randomly selected from the target population, or this was not reported. Only one of the 23 studies reported a sample size justification.[22]

Regarding the statistical analyses, almost a quarter of the papers did not report which analyses were performed, and in only 52% the performed analyses were fully appropriate to answer the research question. The effect of confounders was most frequently corrected in the statistical analyses, and not beforehand in the study design.

In 70% the conclusion of the study was fully supported by the findings, however, in one study the results point to a contrary conclusion than reported.[26]

Explorative analysis

The generalized linear mixed model indicated substantial heterogeneity between studies, meaning that prevalence varies between studies. It was explored whether the overall quality score and the sub score on methodology (questions 1, 4, 7-9, 14-17, 19 in Table 4) were related to the heterogeneity. The goal of this analysis was to check whether selecting studies on quality would narrow the prevalence range substantially. The distance of each study to the median profile in Figure 2 was plotted against the variables of interest. No clear pattern was observed for the quality scores or the sub scores; both low and high quality studies appear on both sides of the median prevalence for all levels. This indicates that the quality of a study did not explain the variation in prevalence, so no studies were excluded for further analyses based on quality score. Furthermore, we investigated whether the heterogeneity was explained by the geographical location (i.e. if the relative difference of a study to the median age-related prevalence fits with an order in geographical location), but no clear trend was visible. For example, the prevalence found by both Bennett[29] and Burke[30] was below the median age-related prevalence curve and the prevalence found by Arafa[36] was above this median, while they all came from the same geographical location: England. On the other hand, prevalences in the Nordic countries all seem to be below the median curve. Instead of trying to understand the influence of geographic location, we calculated, based on our model, 95% prediction limits (the outer

limits in Figure 2). These limits indicate the range of expected true age-related prevalence of DD in observed and unobserved geographical locations in western countries.

Relation between age and prevalence of DD

A combined analysis of 10 studies[12, 15, 22, 25-27, 32, 36-38, 40] representing information on prevalences in different age groups showed an overall relationship that is visualized in the upper graph of Figure 2. In the middle and lower graph of Figure 2, this relationship is shown respectively for females (8 studies[12, 15, 22, 27, 36-38, 40]) and males (11 studies[12, 15, 22, 27, 29, 30, 32, 36-38, 40]). The prevalence is shown as well as the 95% confidence intervals (inner dotted lines), taking into account the heterogeneity between studies. Furthermore, a 95% prediction interval is presented (outer dashed lines), which makes it possible to predict the prevalence at a certain age in a healthy non-hospital population. For instance, the overall prevalence of DD is estimated 12% at 55 years, and 29% at an age of 75 years. The prediction band can be used to estimate the a priori prevalence in a random sample at different ages and geographical locations. Clearly, the prevalence increases with rising age. Furthermore, the graphs show that the prevalence of DD is higher in males than in females. In addition, the age of onset is lower in males compared with the age of onset in females.

Investigating the goodness-of-fit of the estimated models, the R^2 was calculated between the observed numbers of DD, and the predicted numbers of DD from the model. For males the R^2 was estimated at 99.5%, for females the R^2 was equal to 93.0% and for males and females together the R^2 was 97.5%, which demonstrates a good fit of the generalized linear mixed model. This indicates that the models in Figure 2 are able to predict new observations with high certainty. The high goodness of fit may not seem in line with the observed outliers outside the prediction limits in Figure 2. However, several of these outliers were based on small number of subjects (Table 5). For instance, when only one subject is observed in an age category, the prevalence can only be estimated at either 100% of 0% depending on the outcome of DD. The prediction intervals hold true for relative large sample sizes.

.

Discussion

Dupuytren Disease (DD) is a hand disorder, which is often progressive and eventually can cause contractures of the affected fingers. The reported prevalence rates vary widely in the literature. Therefore, the primary goal of this systematic review was to come to a more accurate distribution of the prevalence of DD in the general population. A secondary goal was to perform a meta-analysis on the relation between prevalence of DD and age.

To our knowledge, this systematic review is the first of its kind. First, it focuses on prevalence rates specifically in the general population of western countries, i.e. a healthy non-hospital population, excluding specific patient groups. Second, the quality of the studies was critically assessed. Previous reviews about prevalence of DD concern different kinds of populations, such as manual workers[41], rock climbers[42, 43], and a mixture of healthy participants and patients with a specific disease.[18] Furthermore, geographical location was studied and we performed a thorough meta-analysis on the relationship between age and DD.

Our English search strategy was performed in English databases, so we might have missed relevant articles in foreign languages. However, despite this limitation, several articles in foreign languages, such as German, French and Italian, entered the full text analysis. The kappa for each round of assessment was moderate, emphasizing the necessity of discussing the assessment with multiple authors.

After the full text analysis, 23 studies were included with a number of participants ranging from 37 to 97,537 in the age of 18 to 100 years. Prevalence in these studies varied from 0.6% to 31.6%, which is a smaller range than previously published.[18]

During the quality assessment we came across a number of noteworthy points. First, only few studies reported that they applied sampling to select their participants.[15, 22, 24, 32, 36] However, three of these studies did not describe the method of sampling.[24, 32, 36] If participants are not randomly selected the risk of selection bias increases, which makes it difficult to extrapolate data from these studies. Second, only one study reported a sample size justification.[22] In an observational study, the accuracy of the estimates, i.e. the prevalence, depends on sample size.[44] If a sample size is not calculated on forehand, the results of the study might be less precise than intended. Finally, in only a quarter of the studies the statistical tests were fully stated, and in 52% the analyses were completely appropriate. To enlarge the reproducibility of the

results, it is essential that such information is properly documented. More importantly, to ensure that correct conclusions will be drawn, it is crucial that appropriate analyses are performed.

In order to narrow the prevalence range, we intended to select studies for further analysis, based on their quality. The final overall quality score differed from 0.23 to 0.80. However, in the explorative analysis, no relation was found between this quality score and the reported prevalence. This is in accordance with the findings in a meta-analysis in which the meta-odds ratio for manual work and vibration exposure of all studies was similar to the meta-odds ratio of only high quality studies.[41]

Several articles have been published about the difficulties using an overall score to assess the quality of a study.[45-47] With an overall quality score it is hard to discriminate between poor reporting and poor methodology of the study. Hence, it is advised to evaluate articles based on key components rather then an overall score.[21, 45, 48] Therefore, we analyzed the relation between a high score on methodology and the prevalence of DD. Still, no link was found, so we assumed that the current spread in prevalence was not based on a difference in quality of the studies, but on heterogeneity of the study populations.

We aimed to include studies with participants from a general population. However, we ended with studies that did not provide information about race and that originated mainly from Europe. Nonetheless, the biogeographic regions in Europe differ from Arctic to Mediterranean. Based on our model, we suppose that the prevalence in different geographical locations lies within the prediction interval of Figure 2, but more thorough analyses with additional variables are necessary to clarify and understand the geographic influence on the prevalence of DD.

As mentioned in the results, two studies diagnosed DD differently than other studies.[32, 33] Although this did not change our prevalence range substantially, differences in diagnosing DD complicate the comparison of results. Preferably all stages of DD in all rays are taken into account, for example by using the classification of Iselin or Tubiana.[49, 50] Furthermore, there were differences in reporting age; six studies reported age in categories, without giving the actual range.[26, 27, 34, 34, 36-38] The discrepancies in reporting age also impede comparison of prevalence rates of different studies. Fortunately, we have been able to use data of different age categories in our meta-analysis.

It is well recognized that prevalence of DD increases with rising age, however, until now a thorough analysis on this relationship was lacking. In our meta-analysis, we investigated this relationship by using all

studies that provided information on prevalence in different age categories. We presented the relationship between age and DD, including 95% confidence intervals and 95% prediction intervals.The graphs can be used to determine a common estimate for the prevalence of DD at different ages, both for the total population as well as for males and females separately. Nowadays, still little is known about the prevalence of DD in younger people, since in most studies an age over 50 years was one of the inclusion criteria. However, the relationship between age and prevalence presented in this paper already provides a first indication for prevalence at younger age.

Conclusion

The prevalence of DD in the general population of western countries ranges from 0.6% to 31.6%. With the results of our meta-analysis, we have been able to present the relationship between prevalence of DD and age, including confidence intervals and prediction intervals. With the presented graphs it is possible to determine the prevalence at a certain age for the total general population of western countries, and for men and women separately.

Acknowledgments

The authors would like to thank Professor P.U. Dijkstra, PhD for performing the quality assessment of the study in which some of the authors participated.

References

1. Burge P. Genetics of dupuytren's disease. *Hand Clin*. 1999;15:63-71.

2. Hindocha S, John S, Stanley JK, Watson SJ, Bayat A. The heritability of dupuytren's disease: Familial aggregation and its clinical significance. *J Hand Surg Am*. 2006;31:204-210.

3. Hu FZ, Nystrom A, Ahmed A et al. Mapping of an autosomal dominant gene for dupuytren's contracture to chromosome 16q in a swedish family. *Clin Genet*. 2005;68:424-429.

4. Dolmans GH, Werker PM, Hennies HC et al. Wnt signaling and dupuytren's disease. *N Engl J Med*. 2011;365:307-317.

5. Godtfredsen NS, Lucht H, Prescott E, Sorensen TI, Gronbaek M. A prospective study linked both alcohol and tobacco to dupuytren's disease. *J Clin Epidemiol*. 2004;57:858-863.

6. Bovenzi M. Hand-arm vibration syndrome and dose-response relation for vibration induced white finger among quarry drillers and stonecarvers. italian study group on physical hazards in the stone industry. *Occup Environ Med*. 1994;51:603-611.

7. Savas S, Koroglu BK, Koyuncuoglu HR, Uzar E, Celik H, Tamer NM. The effects of the diabetes related soft tissue hand lesions and the reduced hand strength on functional disability of hand in type 2 diabetic patients. *Diabetes Res Clin Pract*. 2007;77:77-83.

8. Ramchurn N, Mashamba C, Leitch E et al. Upper limb musculoskeletal abnormalities and poor metabolic control in diabetes. *Eur J Intern Med*. 2009;20:718-721.

9. Arafa M, Noble J, Royle SG, Trail IA, Allen J. Dupuytren's and epilepsy revisited. *J Hand Surg Br*. 1992;17:221-224.

10. Beighton P, Valkenburg HA. Bone and joint disorders on tristan da cunha. *S Afr Med J*. 1974;48:743-747.

11. Bergenudd H, Lindgarde F, Nilsson BE. Prevalence of dupuytren's contracture and its correlation with degenerative changes of the hands and feet and with criteria of general health. *J Hand Surg Br*. 1993;18:254-257.

12. Degreef I, De Smet L. A high prevalence of dupuytren's disease in flanders. *Acta Orthop Belg*. 2010;76:316-320.

13. Eadington DW, Patrick AW, Collier A, Frier BM. Limited joint mobility, dupuytren's contracture and retinopathy in type 1 diabetes: Association with cigarette smoking. *Diabet Med*. 1989;6:152-157.

14. French PD, Kitchen VS, Harris JR. Prevalence of dupuytren's contracture in patients infected with HIV. *BMJ*. 1990;301:967.

15. Gudmundsson KG, Arngrimsson R, Sigfusson N, Bjornsson A, Jonsson T. Epidemiology of dupuytren's disease: Clinical, serological, and social assessment. the reykjavik study. *J Clin Epidemiol*. 2000;53:291-296.

16. Geoghegan JM, Forbes J, Clark DI, Smith C, Hubbard R. Dupuytren's disease risk factors. *J Hand Surg Br*. 2004;29:423-426.

17. Critchley EM, Vakil SD, Hayward HW, Owen VM. Dupuytren's disease in epilepsy: Result of prolonged administration of anticonvulsants. *J Neurol Neurosurg Psychiatry*. 1976;39:498-503.

18. Hindocha S, McGrouther DA, Bayat A. Epidemiological evaluation of dupuytren's disease incidence and prevalence rates in relation to etiology. *Hand (N Y)*. 2009;4:256-269.

19. Duin van C, Garssen J. Population forecast 2010-2060: an aging population and increased life expectancy (Bevolkingstrend 2010-2060: sterkere vergrijzing en een langere levensduur). Available at: www.cbs.nl, accessed February 24, 2012.

20. Cho MK, Bero LA. Instruments for assessing the quality of drug studies published in the medical literature. *JAMA*. 1994;272:101-104.

21. Shamliyan T, Kane RL, Dickinson S. A systematic review of tools used to assess the quality of observational studies that examine incidence or prevalence and risk factors for diseases. *J Clin Epidemiol*. 2010;63:1061-1070.

22. Lanting R, Heuvel van den ER, Westerink B, Werker PMN. Prevalence of dupuytren disease in the netherlands. *Plast Reconstr Surg*. 2013;132:394-403.

23. Leemis LM, Trivedi KS. A comparison of approximate interval estimators for the bernoulli parameter. *The American Statistician*. 1996;50:63-68.

24. Ardic F, Soyupek F, Kahraman Y, Yorgancioglu R. The musculoskeletal complications seen in type II diabetics: Predominance of hand involvement. *Clin Rheumatol*. 2003;22:229-233.

25. Noble J, Heathcote JG, Cohen H. Diabetes mellitus in the aetiology of dupuytren's disease. *J Bone Joint Surg Br*. 1984;66:322-325.

26. Noble J, Arafa M, Royle SG, McGeorge G, Crank S. The association between alcohol, hepatic pathology and dupuytren's disease. *J Hand Surg Br*. 1992;17:71-74.

27. Lennox IA, Murali SR, Porter R. A study of the repeatability of the diagnosis of dupuytren's contracture and its prevalence in the grampian region. *J Hand Surg Br*. 1993;18:258-261.

28. Pal B, Griffiths ID, Anderson J, Dick WC. Association of limited joint mobility with dupuytren's contracture in diabetes mellitus. *J Rheumatol*. 1987;14:582-585.

29. Bennett B. Dupuytren's contracture in manual workers. *Br J Ind Med*. 1982;39:98-100.

30. Burke FD, Proud G, Lawson IJ, McGeoch KL, Miles JN. An assessment of the effects of exposure to vibration, smoking, alcohol and diabetes on the prevalence of dupuytren's disease in 97,537 miners. *J Hand Surg Eur Vol*. 2007;32:400-406.

31. Carson J, Clarke C. Dupuytren's contracture in pensioners at the royal hospital chelsea. *J R Coll Physicians Lond*. 1993;27:25-27.

32. Descatha A, Bodin J, Ha C et al. Heavy manual work, exposure to vibration and dupuytren's disease? results of a surveillance program for musculoskeletal disorders. *Occup Environ Med*. 2012;69:296-299.

33. Lucas G, Brichet A, Roquelaure Y, Leclerc A, Descatha A. Dupuytren's disease: Personal factors and occupational exposure. *Am J Ind Med*. 2008;51:9-15.

34. Rafter D, Kenny R, Gilmore M, Walsh CH. Dupuytren's contracture--a survey of a hospital population. *Ir Med J*. 1980;73:227-228.

35. Thomas PR, Clarke D. Vibration white finger and dupuytren's contracture: Are they related?. *Occup Med (Lond)*. 1992;42:155-158.

36. Arafa M, Steingold RF, Noble J. The incidence of dupuytren's disease in patients with rheumatoid arthritis. *J Hand Surg Br*. 1984;9:165-166.

37. Finsen V, Dalen H, Nesheim J. The prevalence of dupuytren's disease among 2 different ethnic groups in northern norway. *J Hand Surg Am*. 2002;27:115-117.

38. Zerajic D, Finsen V. Dupuytren's disease in bosnia and herzegovina. an epidemiological study. *BMC Musculoskelet Disord*. 2004;5:10.

39. Ravid M, Dinai Y, Sohar E. Dupuytren's disease in diabetes mellitus. *Acta Diabetol Lat*. 1977;14:170-174.

40. Mikkelsen OA. The prevalence of dupuytren's disease in norway. A study in a representative population sample of the municipality of haugesund. *Acta Chir Scand*. 1972;138:695-700.

41. Descatha A, Jauffret P, Chastang JF, Roquelaure Y, Leclerc A. Should we consider dupuytren's contracture as work-related? A review and meta-analysis of an old debate. *BMC Musculoskelet Disord*. 2011;12:96-2474-12-96.

42. Straub G. Musculoskeletal problems in rock climbing--a review. *Wien Med Wochenschr*. 2000;150:186-190.

43. Merritt AL, Huang JI. Hand injuries in rock climbing. *J Hand Surg Am*. 2011;36:1859-1861.

44. Biostatistical design of medical studies. In: van Belle G, Fisher LD, Heagerty PJ, Lumley TS, eds. *Biostatistics: A methodology for the health sciences*, Second ed. Hoboken: John Wiley & Sons, Inc.; 2004:19-20.

45. Whiting P, Harbord R, Kleijnen J. No role for quality scores in systematic reviews of diagnostic accuracy studies. *BMC Med Res Methodol*. 2005;5:19.

46. Herbison P, Hay-Smith J, Gillespie WJ. Adjustment of meta-analyses on the basis of quality scores should be abandoned. *J Clin Epidemiol*. 2006;59:1249-1256.

47. Greenland S, O'Rourke K. On the bias produced by quality scores in meta-analysis, and a hierarchical view of proposed solutions. *Biostatistics*. 2001;2:463-471.

48. Soares HP, Daniels S, Kumar A et al. Bad reporting does not mean bad methods for randomised trials: Observational study of randomised controlled trials performed by the radiation therapy oncology group. *BMJ*. 2004;328:22-24.

49. Iselin, M., & Iselin, F. (1967). Maladie de Dupuytren. *Traité de chirurgie de la main;* (pp. 676-678). Brussels: Flammarion.

50. Tubiana R, Michon J, Thomine JM. Scheme for the assessment of deformities in dupuytren's disease. *Surg Clin North Am*. 1968;48:979-984.

51. Attali P, Ink O, Pelletier G et al. Dupuytren's contracture, alcohol consumption, and chronic liver disease. *Arch Intern Med*. 1987;147:1065-1067.

52. Aydeniz A, Gursoy S, Guney E. Which musculoskeletal complications are most frequently seen in type 2 diabetes mellitus?. *J Int Med Res*. 2008;36:505-511.

Figure Legends

Figure 1. Flow-chart of study selection procedure.

Figure 2. Relationship between age and DD. Upper graph: totals, middle graph: females, lower graph: males. Bold line: estimated prevalence, dotted line: 95% confidence interval, dashed line: 95% prediction interval, dots: individual prevalence estimates used in the analysis.

Table Legends

Table 1. Criteria for inclusion and exclusion

DD: Dupuytren Disease

Table 2. Details of included studies.

CG: control group, N: number of participants, SD: standard deviation, CI: confidence interval, IQR: inter quartile range, CAT: age reported only in categories.

Table 3. Quality assessment of included studies per question.

n: number of studies, %: percentage, NA: not applicable, † See Table 3, ‡ Open question which does not contribute to final score, * Case studies were not included, so question 6 was not applicable for each of the included articles, ** Questions were not applicable, because this concerns intervention studies.

Table 4. Quality assessment of included studies per study.

† Questions: **1**: Study design, **2**: Research question, **3**: Study question sufficiently described, **4**: Study design appropriate to answer study question, **5**: Inclusion and exclusion criteria specified, **6**: Case studies: patient characteristics adequately reported, **7**: Subjects appropriate to study question, **8**: Control subjects appropriate, **9**: Random selection of subjects, **10**: Method of random selection sufficiently well described, **11**: Random allocation to treatment group sufficiently described, **12**: Blinding of investigators to intervention reported, **13**: Blinding of subjects to intervention reported, **14**: Measurement bias accounted for by methods other than blinding, **15**: Known confounders

accounted for by study design, **16**: Known confounders accounted for by analysis, **17**: Sample size justification, **18**: Post hoc power calculations or confidence intervals reported for statistically non significant results, **19**: Appropriate statistical analyses, **20**: Statement of statistical tests, **21**: Exact values of confidence intervals reported for each test, **22**: Reporting of attrition of subject and reason for attrition, **23**: Results completely reported for subjects who completed the study, **24**: Findings support the conclusion.

Question 1 was scored 3 (cohort design) or 2 (cross-sectional design), other questions were scored 2 (yes), 1 (partial), 0 (no), NA (not applicable).

The score was calculated by dividing the total points by the maximum possible points. A higher score represents a higher quality.

‡ Open question which does not contribute to the final score.

* Case studies were not included, so question 6 was not applicable for each of the included articles.

** Questions were not applicable, because this concerns intervention studies.

Table 5. Studies outside prediction intervals

Age cat: age category, n DD: participants with DD, n total: total participants, % DD: percentage of participants with DD, 95% PI: 95% prediction interval

† Outlier not visible in Figure 2 (Y-axis ranges from 0-80%)

Textbox 1. Criteria for inclusion and exclusion

Round 1. Title and abstract *Inclusion criteria:* - DD as research theme - General population as sample *Exclusion criteria:* - Case report - Case series - Review article - Subjects aged <18 years
Round 2. First full text assessment *Inclusion criteria:* - Prevalence of DD as research theme *Exclusion criteria:* - Age is not reported - Physical examination to diagnose DD was not performed or not reported - Full text is not available
Round 3. Second full text assessment *Inclusion criteria:* - Prevalence is calculated - Data is provided to calculate prevalence *Exclusion criteria:* - Unclear how DD is diagnosed - Outcome is 'Dupuytren Contracture', not further specified - Incidence was reported instead of prevalence

DD: Dupuytren Disease

Table 1. Details of included studies.

Authors	Year	Population	N	Gender	Age			Prevalence (95%CI)	Quality score
					Mean	SD	Range		
Arafa[36]	1984	Patients of fracture clinic (CG)	555	F and M	CAT			16.0 [13.1 ; 19.4]	0.46
Ardic[24]	2003	Non-diabetic patients of department of physical medicine and rehabilitation (division rheumatology) (CG)	37	F and M	55.7	11.5	30-79	2.7 [1.0 ; 14.2]	0.44
Attali[51]	1987	Patients of gastroenterology unit without alcoholism or chronic liver disease (CG)	174	F and M	58.9	22.7		12.5 [8.1 ; 18.5]	0.49
Aydeniz[52]	2008	Non-diabetic patients of public health clinic (CG)	101	F and M	60.1	7.6		4.0 [1.1 ; 9.8]	0.51
Bennettt[29]	1982	Workers PVC manufacturing plant not involved with bagging or packing (CG)	84	M	40.1			1.19 [0.0 ; 6.5]	0.46
Burke[30]	2007	Miners seeking compensation for Hand-Arm Vibration Syndrome	97537	M	53.5			8.13 [8.0 ; 8.30]	0.62
Carson[31]	1993	Ex-military service pensioners in the Royal Hospital Chelsea	400	M	75.9		65-99	13.8 [10.5 ; 17.5]	0.38
Degreef[12]	2010	Visitors of markets in Flanders, Belgium	500	F and M	70.4		50-100	31.6 [27.5 ; 35.9]	0.46
Descatha[32]	2012	Employees in private sector in Pays de la Loire, France	2161	M	38.5		20-59	1.25 [0.8 ; 1.8]	0.66
Eadington[13]	1989	Normotensive, non-diabetic subjects, selected from inpatients, outpatients and hospital staff members (CG)	150	F and M	51.2	17.4		18.0 [12.2 ; 25.1]	0.64
Finsen[37]	2002	Residents of rural municipalities in Norway	456	F and M	CAT		50-80+	7.5 [5.05 ; 9.87]	0.51
Gudmundsson[15]	2000	Residents of Reykjavik and adjacent communes, Iceland	2165	F and M	57.5		45-94	13.3 [11.9 ; 14. 8]	0.56

Lanting[22]	2013	Residents of Groningen, The Netherlands	763	F and M	62 (median)		56-69 (IQR)	22.1 [19.3 ; 25.3]	0.80
Lennox[27]	1993	Patients on geriatric ward, not admitted for hand pathology	200	F and M	CAT			30.0 [23.7 ; 36.9]	0.37
Lucas[33]	2008	Civil servants of Pays de la Loire and Brittany, France	2406	M	45.3	7.6		8.8 [7.7 ; 10.0]	0.64
Mikkelsen[40]	1972	Residents of Haugesund, Norway	15950	F and M	45.0		16-99	5.6 [5.3 ; 6.0]	0.46
Noble[26]	1992	Patients of fracture clinic (CG)	100	F and M	CAT	.		8.0 [3.5 ; 15.2]	0.36
Noble[25]	1984	Patients of fracture clinic (CG)	150	F and M	57.4			18.0 [12.2 ; 25.1]	0.28
Pal[28]	1987	Non-diabetic subjects without musculoskeletal complaints (CG)	75	F and M	44.0 (median)		18-76	9.0 [3.8 ; 18.3]	0.49
Rafter[34]	1980	Inpatients in acute medical and surgical wards	403	M	CAT	.		17.1 [13.6 ; 21.2]	0.23
Ravid[39]	1977	Non-diabetic patients of different departments of medicine (CG)	1396	F and M	52.0		19-86	0.6 [0.3 ; 1.2]	0.49
Thomas[35]	1992	Patients admitted to general surgical ward (CG)	150	M	64.1		50-85	10.7 [6.2 ; 16.7]	0.46
Zerajic[38]	2004	Visitors of public places in both urban and rural areas of Bosnia Herzegovina	1207	F and M	CAT	.		25.4 [23.0 ; 28.0]	0.59

CG: control group, N: number of participants, SD: standard deviation, CI: confidence interval, IQR: inter quartile range, CAT: age reported only in categories.

Table 2. Quality assessment of included studies per question.

	Answer:	Yes		Partial		No		NA	
	Question	n	%	n	%	n	%	n	%
1	Study design †								
2	What was the study question? ‡								
3	Was the study question sufficiently described?	5	22%	15	65%	3	13%	0	0%
4	Was the study design appropriate to answer the study question?	21	91%	2	9%	0	0%	0	0%
5	Were both inclusion and exclusion criteria specified?	3	13%	6	26%	14	61%	0	0%
6	For case studies only: Were patient characteristics adequately reported?*	0	0%	0	0%	0	0%	23	100%
7	Were subjects appropriate to the study question?	19	83%	4	17%	0	0%	0	0%
8	Were control subjects appropriate?	12	52%	6	26%	5	22%	0	0%
9	Were subjects randomly selected from the target population?	5	22%	0	0%	18	78%	0	0%
10	If subjects were randomly selected, was the method of random selection sufficiently well described?	1	4%	1	4%	3	13%	18	78%
11	If subjects were randomly allocated to treatment groups, was method of random allocation sufficiently described?**	0	0%	0	0%	0	0%	23	100%
12	If blinding of investigators was possible, was it reported?**	0	0%	0	0%	0	0%	23	100%
13	If blinding of subjects to intervention was possible, was it reported?**	0	0%	0	0%	0	0%	23	100%
14	Was measurement bias accounted for by other methods than blinding?	6	26%	11	48%	6	26%	0	0%
15	Were known confounders accounted for by study design?	5	22%	3	13%	13	57%	2	9%
16	Were known confounders accounted for by	9	39%	5	22%	7	30%	2	9%

	analysis?								
17	Was there a sample size justification before the study?	1	4%	0	0%	22	96%	0	0%
18	Were post hoc power calculations or confidence intervals reported for statistical non significant results?	4	17%	4	17%	15	65%	0	0%
19	Were statistical analyses appropriate?	12	52%	5	22%	6	26%	0	0%
20	Were the statistical tests stated?	6	26%	12	52%	5	22%	0	0%
21	Were exact values or confidence intervals reported for each test?	5	22%	13	57%	5	22%	0	0%
22	Were attrition of subjects and reason for attrition recorded?	4	17%	3	13%	16	70%	0	0%
23	For those subjects who completed the study; were results completely reported?	15	65%	7	30%	1	4%	0	0%
24	Do the findings support the conclusions?	16	70%	6	26%	1	4%	0	0%

n: number of studies, %: percentage, NA: not applicable, † See Table 3, ‡ Open question which does not contribute to final score, * Case studies were not included, so question 6 was not applicable for each of the included articles, ** Questions were not applicable, because this concerns intervention studies.

Table 3. Quality assessment of included studies per study

Author	Questions†																									
	1	2‡	3	4	5	6*	7	8	9	10	11**	12**	13**	14	15	16	17	18	19	20	21	22	23	24	Total	Max. points	Score
Arafa[36]	2		0	1	0	NA	2	1	2	0	NA	NA	NA	2	2	1	0	0	0	1	1	0	2	2	19	41	0.46
Ardic[24]	2		1	2	0	NA	2	0	2	0	NA	NA	NA	1	0	2	0	0	1	1	0	0	2	2	18	41	0.44
Attali[51]	2		1	2	0	NA	2	2	0	NA	NA	NA	NA	0	0	2	0	0	2	1	1	0	2	2	19	39	0.49
Aydeniz[52]	2		1	2	1	NA	2	2	0	NA	NA	NA	NA	2	2	0	0	0	1	1	1	0	2	1	20	39	0.51
Bennett[29]	2		1	2	0	NA	2	2	0	NA	NA	NA	NA	1	0	2	0	0	1	1	0	0	2	2	18	39	0.46
Burke[30]	2		2	2	0	NA	2	2	0	NA	NA	NA	NA	1	0	2	0	2	2	2	2	0	1	2	24	39	0.62
Carson[31]	2		0	2	0	NA	2	2	0	NA	NA	NA	NA	1	0	1	0	0	0	0	1	0	2	2	15	39	0.38
Degreef[12]	2		1	2	2	NA	2	2	0	NA	NA	NA	NA	2	0	0	0	0	0	0	1	0	2	2	18	39	0.46
Descatha[32]	2		1	2	2	NA	2	2	2	0	NA	NA	NA	1	0	2	0	2	2	1	1	1	2	2	27	41	0.66
Eadington[13]	2		2	2	2	NA	2	1	0	NA	NA	NA	NA	1	1	2	0	1	2	2	1	1	2	1	25	39	0.64
Finsen[37]	2		1	2	1	NA	2	0	0	NA	NA	NA	NA	0	0	1	0	1	2	2	2	2	1	1	20	39	0.51
Gudmundsson[15]	3		1	2	0	NA	2	2	2	1	NA	NA	NA	1	0	2	0	1	2	1	1	0	1	1	23	41	0.56
Lanting[22]	2		2	2	0	NA	2	2	2	2	NA	NA	NA	1	2	2	2	2	2	2	2	0	2	2	33	41	0.80
Lennox [27]	2		1	2	0	NA	1	0	0	NA	NA	NA	NA	1	NA	NA	0	0	1	1	1	0	1	2	13	37	0.37
Lucas[33]	2		1	2	1	NA	1	2	0	NA	NA	NA	NA	1	0	2	0	2	2	2	2	1	2	2	25	39	0.64
Mikkelsen[40]	2		1	2	1	NA	2	0	0	NA	NA	NA	NA	2	NA	NA	0	0	0	0	0	2	2	2	16	35	0.46
Noble [25]	2		0	2	0	NA	2	1	0	NA	NA	NA	NA	0	2	0	0	0	0	0	1	0	0	1	11	39	0.28
Noble[26]	2		1	2	0	NA	1	1	0	NA	NA	NA	NA	0	1	0	0	0	2	1	1	0	2	0	14	39	0.36
Pal[28]	2		2	2	1	NA	2	2	0	NA	NA	NA	NA	2	0	0	0	0	1	1	1	2	1	0	19	39	0.49
Rafter[34]	2		1	1	0	NA	1	2	0	NA	NA	NA	NA	0	0	0	0	0	0	0	0	0	1	1	9	39	0.23
Ravid[39]	2		2	2	0	NA	2	1	0	NA	NA	NA	NA	0	2	1	0	0	2	1	0	0	2	2	19	39	0.49
Thomas[35]	2		1	2	0	NA	2	1	0	NA	NA	NA	NA	1	0	1	0	0	2	1	2	0	1	2	18	39	0.46

Zerajic[38]	2	1	2	1	NA	2	0	0	NA	NA	NA	NA	2	1	0	0	1	2	2	1	2	2	2	23	39	0.59

† Questions: **1**: Study design, **2**: Research question, **3**: Study question sufficiently described, **4**: Study design appropriate to answer study question, **5**: Inclusion and exclusion criteria specified, **6**: Case studies: patient characteristics adequately reported, **7**: Subjects appropriate to study question, **8**: Control subjects appropriate, **9**: Random selection of subjects, **10**: Method of random selection sufficiently well described, **11**: Random allocation to treatment group sufficiently described, **12**: Blinding of investigators to intervention reported, **13**: Blinding of subjects to intervention reported, **14**: Measurement bias accounted for by methods other than blinding, **15**: Known confounders accounted for by study design, **16**: Known confounders accounted for by analysis, **17**: Sample size justification, **18**: Post hoc power calculations or confidence intervals reported for statistically non significant results, **19**: Appropriate statistical analyses, **20**: Statement of statistical tests, **21**: Exact values of confidence intervals reported for each test, **22**: Reporting of attrition of subject and reason for attrition, **23**: Results completely reported for subjects who completed the study, **24**: Findings support the conclusion.

Question 1 was scored 3 (cohort design) or 2 (cross-sectional design), other questions were scored 2 (yes), 1 (partial), 0 (no), NA (not applicable).

The score was calculated by dividing the total points by the maximum possible points. A higher score represents a higher quality.

‡ Open question which does not contribute to the final score.

* Case studies were not included, so question 6 was not applicable for each of the included articles.

** Questions were not applicable, because this concerns intervention studies.

Table 4. Studies outside prediction intervals

Population	Age cat.	Author	n DD	n total	% DD	95% PI
Total	<30	Arafa[36]	1	34	2.94	0.02 – 0.61
	30-34	Mikkelsen[40]	1	1043	0.10	0.12 – 2.89
	30-39	Arafa[36]	4	47	8.51	0.18 – 4.38
	30-39	Noble (1984)[25]	1	5	20	0.18 – 4.38
	50-59	Finsen[37]	2	103	1.94	2.53 – 27.46
	61-65	Degreef[12]	32	86	37.21	5.25 – 37.20
	75-79	Zerajic[38]	43	72	59.72	12.26 – 49.43
	76-80	Lanting[22]	30	57	52.63	12.84 – 50.15
	>80	Finsen[37]	0	24	0	16.76 – 54.41
	95-99	Mikkelsen[40]	0	3	0	24.32 – 60.98
Males	<30	Descatha[32]	0	491	0	0.06 – 4.22
	55-64	Bennett[29]	0	9	0	2.84 – 42.03
	75-79	Zerajic[38]	30	40	75	7.60 – 53.80
	76-80	Lanting[22]	18	24	75	7.95 – 54.34
	>80	Finsen[37]	0	7	0	10.38 – 57.52
	80+	Zerajic[38]	24	40	60	9.41 – 56.35
	81-85	Lanting[22]	8	14	57.14	9.80 – 56.83
	>90	Lennox[27]	4	6	66.67	14.52 – 61.62
	90-94	Mikkelsen[40]	1	1	100†	13.45 – 60.68
	95-99	Burke[30]	0	1	0	15.60 – 62.53
	95-99	Mikkelsen[40]	0	1	0	15.60 – 62.53
Females	81-85	Lanting[22]	8	17	47.06	0.25 – 46.83

Age cat: age category, n DD: participants with DD, n total: total participants, % DD: percentage of participants with DD, 95% PI: 95% prediction interval

† Outlier not visible in Figure 2 (Y-axis ranges from 0-80%)