Content Scraping Developer at Monocl – Case Assignment

Overall Objective: To transform unstructured information about participants at scientific conferences into structured data about the persons and their role and presentation at the conference. This structured data will then be implemented in Monocl's SaaS-platform (www.monocl.com) and be used by primarily users in the pharmaceutical industry who use it to identify and evaluate which scientific researchers they should be working with in different collaboration-projects. The information about how often the researchers are out presenting on scientific conferences is often a key factor that the users of the platform use to evaluate which of these researchers are influential enough to work with.

Case Assignment: You will in this Case Assignment gather parts of the data from a scientific conference called "5th World Psoriasis and Psoriatic Arthritis Conference 2018" held in Stockholm earlier this year (http://www.ifpaworldconference.com/. What we will be using for this assignment is however only the Abstract Book which is attached in this email.

In the pdf-document there is first a couple of pages with introductory material which you can ignore and then you will start to see a lot of pages looking like the screenshot below. In red you will find the Topic Title and in green you find the information about the persons (their name, their affiliation/organization and in some cases their geographical location) and in yellow you find the presentation abstract. This structure repeat itself over many pages and **your task is to find a way of extracting the data** that is presented in this repetitive structure.

≥10%) was based on the respondent's estimate of palms of affected skin (1 palm = 1% BSA).
Results: Among 22,050 individuals questioned, reported prevalence of PsO and/or PsA was 9.7% (9.4% in Sweden, 9.2% in Denmark, and 11.9% in Norway); 1264 (5.7%) individuals reported physician diagnosis and 1221 of these completed the full survey. Among survey participants, 74.6% reported PsO alone, 10.3% PsA alone, and 15.1% both. There was a limited correlation between self-perceived disease severity and estimates of % BSA for PsO (Spearman's rho 0.42). Self-perceived PsO severity was related to the most common symptoms (itching and flaking/scales) but % BSA-based clinical severity was not. Respondents with PsA+PsO reported significantly worse self-perceived severity and % BSA of PsO than those with PsO alone. Overall, PsA symptoms were self-perceived to be more severe than PsO symptoms.
Conclusion: Patient perceptions of PsO severity differ from their estimates of %BSA and should be considered in the assessment and management of PsO.

P077

PATIENT EXPECTATIONS AND SATISFACTION IN PSORIASIS TREATMENT: A SURVEY FROM EUROPE AND CANADA

Jonathan Barker[1], Robert McKenzie[2], Daniel Saure[2], Stefan Wilhelm[2], Kunal Gulati[2], Ronald Vender[1]
[1]Guy's Hospital, St. John's Institute of Dermatology, London, UK, [2]Eli Lilly and Company, [3]Dermatrials Research INC, Hamilton, Ontario, Canada

Introduction: Psoriasis is a chronic immune-mediated disease that negatively affects the quality of life of patients.1 Despite recent advancements in treatment, patient expectations may differ from the treatment outcome, potentially leading to dissatisfaction.
Objective: To evaluate patient expectations on psoriasis treatment outcomes and to what extent these were met by the treatment.
Methods: Data were collected from patients with moderate-to-severe psoriasis with ≥ 3% skin involvement. Patients were included from 17 European countries and Canada through a structured, web-based interview in their local language. Patients were on topical, PUVA/phototherapy, conventional systemic or biological treatment. Descriptive statistical analyses were conducted.
Results: Overall, 1946 patients (mean [SD] age: 42 [13] years, 43% male) with psoriasis were surveyed between 6 July 2016 and 5 May 2017. The vast majority of patients currently on treatment reported reduced itchiness (69%), reduced flaking (62%), and clearer skin/skin that looked better generally (61%) as the main expectations of their treatment. In addition, these patients reported duration of effect (41%), extent of overall skin clearance (36%), and speed of onset (34%) as the main influences for changing their current treatment. Furthermore, patients were asked whether their achieved results matched with their initial treatment expectations. Patients reported that the convenience of dosing (68%) and administration method (68%) completely met, or even exceeded their initial expectations for the treatment. Conversely, expectations for reduced itchiness (62%), reduced flaking (62%), clearer skin/ skin that looked better generally (63%) were only partially met or not met at all. Approximately, two-third of patients were dissatisfied or only partially satisfied with the extent of their skin clearance and improvement in at least one specific area. Specifically, scalp (48%), legs (42%), and arms (41%) were the highest rated areas if asked for insufficient satisfaction for clearance.
Conclusion: Data from this international survey indicate that efficacy and maintenance of response are the patients' primary expectations of psoriasis treatment whereas, insufficient response to therapy in difficult-to-treat areas appears to be a major reason for patient dissatisfaction.
Reference:
1. Paul C, Bang B, Lebwohl M. Fixed combination calcipotriol plus betamethasone dipropionate aerosol foam in the treatment of psoriasis vulgaris: rationale for development and clinical profile. Expert Opin Pharmacother. 2017 Jan;18(1):115-121.
This survey was sponsored by Eli Lilly and Company, Indianapolis, USA.

P078

SOCIOECONOMIC DETERMINANTS OF PAEDIATRIC PSORIASIS

Jonathan Groot[1], Anne-Marie Nybo Andersen[2], Abdulfatah Adam[2], Tilde Elkjær Tind Nielsen[2,3], Christoffer Blegvad[3], Lone Skov[3]
[1]Department of Dermatology and Allergy, Herlev and Gentofte Hospital, University of Copenhagen, Hellerup, Denmark, [2]Section for Epidemiology, Department of Public Health, University of Copenhagen, Copenhagen, Department of Dermatology and Allergy, Herlev and Gentofte Hospital, University of Copenhagen, Hellerup, Denmark

Introduction: Psoriatic comorbidities have been shown to be inversely associated with socioeconomic position (SEP). We hypothesized that the social patterning of paediatric psoriasis would be similar to that observed for several psoriatic comorbidities, such as cardiovascular disease and type 2 diabetes mellitus.
Objectives: Our aim was to investigate whether maternal SEP is a determinant of paediatric psoriasis.
Methods: Data on paediatric psoriasis from 36,003 offspring of a national birth cohort were cross-linked with nation-wide registry data on maternal educational attainment, maternal labour market attachment and equivalised household income. We performed logistic regression analyses to test for associations between measures of maternal SEP and paediatric psoriasis. Cohort analyses were conducted, estimating the odds ratios (OR) and 95% confidence intervals (CI) of offspring psoriasis. We included maternal age at birth and maternal psoriasis as covariates in adjusted analyses.
Results: Maternal educational attainment and equivalised household income were inversely associated with offspring psoriasis. Offspring of mothers with a low educational attainment had an OR 1.62 (95% CI: 1.20–2.18; $p < 0.01$) of paediatric psoriasis, after adjusting for maternal psoriasis and age, compared to offspring of mothers with a medium educational attainment. Offspring of mothers in the highest quartile compared to mothers in the lowest quartile of equivalised maternal household income had an OR 0.59 (95% CI: 0.44–0.80; $p < 0.001$) of paediatric psoriasis, after adjusting for maternal psoriasis and age.
Conclusion: Lifetime prevalence of paediatric psoriasis was inversely associated with maternal SEP in our study population. Early life exposure to modifiable risk factors associated with SEP may play an important role in the development of paediatric psoriasis. Future studies are warranted to clarify the role of mediating factors.

P079

OBESITY AND THE RISK OF PSORIASIS: A KOREAN NATIONWIDE, POPULATION-BASED STUDY

Ju Hee Han[1], Kyung Do Han[2], Chul Hwan Bang[1], Young Min Park[1], Jun Young Lee[1], Yong Gyu Park[2], Ji Hyun Lee[1]
Department of Dermatology, Seoul St. Mary's Hospital, College of Medicine, The Catholic University of Korea, Department of Biostatistics, College of Medicine, The Catholic University of Korea, Seoul, Korea

Introduction: Psoriasis is a T-helper-1 and -17 cell-mediated, chronic, inflammatory skin disease affecting approximately 1–3% of the general population. Psoriasis is currently considered as a chronic low-grade inflammatory condition that plays an active role in the development of the pathophysiologic phenomena responsible for metabolic syndrome and cardiovascular disease. Obesity, as a component of the metabolic syndrome, represents a major comorbidity and possibly contributes to reduced treatment response.
Objective: This nationwide, population-based study investigated the impact of body mass index (BMI) and waist circumference (WC) on the psoriasis in the Korean population.
Methods: We used the health check-up database, which is sub-

Data is already filled in for the presentations P098 and P099 from page 41 in the abstract book and your task is to continue for the next couple of pages and find a way of extracting the information and then fill it in the Excel- document. We understand that you might not be able to extract all the data separated in to different fields and that it

might require some final manual steps. But to as great extent as possible we want the data extracting to be done through coding/scraping.

Final remarks: We are well aware of that the instructions are limited for this case and there is no exact way of performing this task as the data can be interpreted in different ways, but we want to see how you take on this assignment with limited information. When this is done for real later on, you will have more instructions and colleagues to discuss around how to map the data. As a part of this assignment we will afterwards sit down together with you and discuss your approach to this problem and look at the solution you ended up using for the problem at hand.