Contents

1 Introduction

Psoriasis is a chronic inflammatory disease of the skin with a strong genetic component, affecting roughly 2% of the global population. It often presents as red, scaly plaques that can appear on the extensor surfaces of the arms and legs (e.g., elbows and knees), the trunk, and the scalp. These plaques are the result of an accelerated turnover of skin cells (keratinocytes) and an aberrant inflammatory response driven predominantly by T lymphocytes. Although psoriasis is most visible on the skin, recent insights suggest it is a systemic condition rather than a disease confined solely to the epidermis. Elevated levels of pro-inflammatory cytokines—such as TNF-alpha, IL-17, and IL-1beta—are commonly found in patients, indicating a state of chronic inflammation that can also involve the joints (leading to psoriatic arthritis) and the cardiovascular system.

Despite being a relatively common disease, psoriasis diagnosis remains largely clinical, based on visual examination of the plaques and patient history, due to the lack of universally accepted biomarkers. Histological examination of affected skin typically shows thickening of the epidermis, reduced or absent granular layer, dilated blood vessels in the superficial dermis, and a marked infiltration of immune cells. Because it significantly affects both physical appearance and general health, psoriasis has considerable social and medical impact, emphasizing the need for ongoing research into its underlying mechanisms and novel treatment approaches. Furthermore, recent biological and clinical studies have uncovered a subtle but deep relation with conditions of male infertility in different psoriatic patients. This is precisely, the point that this study tries to focus on.

According to the World Health Organization (WHO), infertility is defined as the absence of conception after 12–24 months of regular, unprotected intercourse. However, the term "infertility" can be refined into true "sterility," where no medical intervention can restore fertility, and "subfertility," which may be overcome with appropriate treatments. About 15–20% of couples experience difficulty conceiving, influenced partly by social factors (e.g., delayed family planning).
Male infertility can be classified as pre-testicular (secretory) or testicular

(secretory) in origin [1]:

- Pre-testicular (secretory) causes involve hypothalamic-pituitary dysfunction (e.g., Kallmann syndrome), often leading to low gonadotropins, low testosterone, and azoospermia.
- Testicular (secretory) causes include various disorders affecting the seminiferous tubules. They can stem from:
  - Congenital factors, such as Klinefelter syndrome (XXY), Y chromosome microdeletions (e.g., AZF regions), or cryptorchidism (undescended testes).
  - Acquired factors, like orchitis, torsions, tumors (and subsequent chemotherapy or radiotherapy), varicocele, thyroid or prolactin abnormalities, and autoimmune processes against sperm cells.

The testis is considered an immunologically privileged site, protected from systemic inflammation by tight junctions between Sertoli cells and the blood–testis barrier. While spermatogenesis occurs inside the testis, seminal fluid is produced in the prostate and seminal vesicles outside this protected environment. Consequently, systemic inflammatory responses—driven by cytokines such as IL-17, TNF-alpha, and IL-6—can still affect sperm quality and function.
Among these cytokines, IL-6 is secreted by various testicular cell types (including Sertoli and Leydig cells) and helps regulate processes such as transferrin synthesis in Sertoli cells and the progression of germ cells through meiosis. However, excessive IL-6 signaling can disrupt steroidogenesis in Leydig cells, leading to reduced testosterone production. IL-17 can similarly inhibit gonadotropin signaling, and TNF-alpha has been implicated in modulating steroidogenesis and promoting apoptosis of abnormal sperm.
Conversely, IL-10 appears to be the main anti-inflammatory cytokine in humans; higher IL-10 levels correlate with fertility, possibly due to its ability to inhibit immune recognition of sperm antigens. Elevated IL-10 in conditions like cryptorchidism may prevent an autoimmune response against the retained testis. These observations suggest that measuring pro- and anti-inflammatory cytokines in seminal fluid could be a valuable addition to

[1] Naz, M., Kamal, M. (2017). Classification, causes, diagnosis, and treat- ment of male infertility: A review. Oriental Pharmacy and Experimental Medicine, 17, 89–109.

routine screening for male infertility. Additionally, psoriasis, as a systemic inflammatory disease, may influence spermatogenesis and thus be considered among the testicular causes of male infertility.

In a comprehensive study, **Damiani (2023)** investigated the impact of IL-23 inhibitors on male fertility in patients with moderate-to-severe psoriasis. First, untargeted metabolomics compared psoriatic skin (lesional and non-lesional) with that of healthy controls, revealing significant dysregulation ($p<0.001$) of metabolites such as L-acetylcarnitine and hypotaurine—both of which have been previously linked to male infertility. A subsequent systematic literature review, which also assessed Italian clinical guidelines published since 1990, showed a notable absence of data regarding sperm function in psoriatic patients. These findings were supported by a survey of Italian assisted reproductive technology centers, which likewise highlighted a general lack of awareness about this potential comorbidity.
Damiani then performed a multi-pronged analysis to solidify the link between psoriasis and male infertility. Data from the Global Burden of Diseases (GBD) database and the CLALIT Israeli healthcare system indicated that individuals with psoriasis often demonstrated unsatisfactory sperm DNA fragmentation indices (DFI).
IL-23 inhibitors, more than other psoriasis treatments, can prevent PsA onset, Damiani scrutinized the crystallographic properties of epitopes targeted by various anti–IL-23 agents. Of the 572 p19 subunit variants noted in a genome-wide association study (GWAS), 10 non-synonymous polymorphisms specifically mapped onto these epitopes and conveyed some level of drug resistance. Assessment of both therapeutic efficacy and possible drug-induced erectile dysfunction pinpointed risankizumab as having the most favorable profile. Notably, 11 psoriasis patients with DFI $> 30\%$ showed a statistically significant improvement in sperm DNA integrity at 16 weeks of risankizumab therapy.
Lastly, the study proposed a dedicated diagnostic and therapeutic care pathway (PDTA) for men of reproductive age with moderate-to-severe psoriasis. This PDTA underscores the value of early and continuous IL-23 inhibitor use to minimize cumulative inflammatory damage, preserve sperm quality, and potentially enhance the success of medically assisted reproduction.

The goal of this thesis is to investigate the complex relationship between psoriasis and male infertility, focusing on understanding whether specific co-

morbidities contribute to or result from the coexistence of these conditions. The research adopts a multi-faceted approach, combining statistical analysis, clustering methods, and predictive modeling to identify patterns and associations within patient cohorts.
The scope of the study includes the following key areas:

- **Data Collection and Cohort Definition:** The dataset, sourced from the CLALIT Health Services Israeli database, includes 280,450 patients equally divided between psoriasis and control groups. The dataset captures clinical diagnoses, comorbidities, and temporal relationships between psoriasis and infertility.
- **Comorbidity Analysis:** Statistical tests, including chi-squared tests, were performed to identify comorbidities significantly associated with psoriasis and infertility. This analysis aimed to uncover shared risk factors and systemic conditions linked to both diseases.
- **Chronological Analysis:** Two distinct pathways were analyzed: patients developing infertility after psoriasis and those developing psoriasis after infertility. This approach helped clarify the temporal relationship between the conditions.
- **Clustering and Classification:** Machine learning techniques, including Gaussian Mixture Models (GMM) and the StepMix latent class model, were employed to identify distinct patient subgroups and predict infertility outcomes based on psoriasis-related features.
- **Comorbidome Construction:** Logistic regression models were used to compute odds ratios (ORs) for comorbidities associated with infertility among psoriatic patients, enabling the construction of a comprehensive comorbidome.
- **Matched Cohort Analysis:** To address potential biases, age-based matching was applied, ensuring balanced comparisons between infertile and non-infertile psoriasis patients.

The study's scope extends beyond establishing correlations, aiming to identify potential causal pathways and risk profiles. By combining epidemiological insights with advanced analytical techniques, this research provides a foundation for personalized treatment strategies and future clinical investigations.

2 Data and Methods

2.1 Data Description

The following work is based on the analysis of a dataset retrieved from the CLALIT Health Services Israeli database. This is the largest provider of public and semi-private health services in Israel. Under Israeli law, it is run as a not-for-profit entity. The data at hand has been collected between January, 2002 and December 2022. It has been optimally designed to study the relation between Psoriasis and Male Infertility in the available subjects. Full anonimity of subjects has been assured by the Services of Israel, ensuring all patient information was de-identified.

The CLALIT dataset, maintained by Israel's largest health services organization, provides comprehensive records for randomly selected members who receive routine healthcare services. All medical diagnoses (including psoriasis and various comorbidities) are systematically entered by the treating physicians following standardized classification protocols. For the purposes of this study, the sample was designed to be balanced: half of the individuals carried a clinical diagnosis of psoriasis, while the other half did not. Moreover, to maintain an equitable representation across genders, the final cohort was split evenly between males and females (50% each). This approach ensured a robust comparison between psoriasis and non-psoriasis groups while minimizing potential bias stemming from gender imbalances. Additionally, the dataset exhibits a pronounced imbalance in the proportion of male subjects diagnosed with infertility—around 5% among all male participants and approximately 2% of the entire cohort. This skew arises from the dataset's original design, which prioritized capturing a Psoriasis group and a corresponding Control group. Consequently, while other comorbidities (including infertility) are recorded for each patient, they were not the primary focus of the data collection process.

The dataset at hand consists of 280,450 patients, equally split between a Psoriasis group and a Control group. Starting from the original CLALIT dataset, for each patient we recorded the (i) date of birth, (ii) group in which the patient ended (Psoriasis versus Control), (iii) age, (iv) sex, (v) the Psoriasis diagnosis date, which would be unavailable under Control group, and a series of comorbidities with the potential date of diagnosis. Infertility is

one of these comorbidities. They also include different forms of malignancies, Tuberculosis, Syphilis, Gonorrhea, Hepatitis B and C, Obesity, Diabetes, Hyperlipidemia, Wilson Disease, Gaucher Disease, thus a total of 122 effective comorbidities.

Within the overall dataset, a total of 69,342 male patients (24.72% of the entire cohort) were allocated to the Psoriasis group, and the same number (69,342; 24.72%) formed the Control group. Of all individuals in the dataset, 14,528 (5.18%) held an official diagnosis of psoriasis. Among the male population, 802 (0.29%) were documented as infertile, with their distribution evenly split: 406 (0.14% of the total dataset) fell under the Psoriasis group and 396 (0.14%) under the Control group. Notably, of those 406 infertile men in the Psoriasis group, 59 (0.02%) received the infertility diagnosis prior to the onset of psoriasis, whereas 40 (0.014%) were diagnosed afterward. These figures underscore both the balanced design of the dataset—in terms of group membership—and the relatively low prevalence of diagnosed infertility in the male population, illuminating an important demographic and clinical baseline for subsequent analyses.

2.2 Age Distribution

In terms of **Age Distribution**, the minimum age recorded is 0, while the maximum is 103; the mean age in the sample is 42.987 and the median is 43. The standard deviation is 20.978. We can go one step further in the demographics of age and compare distributions for Psoriasis and Control groups.
The average age in the Psoriasis group is 42.978 and the median is 43; the minimum age is 0 and the maximum is 103. The standard deviation is 20.971. On the other hand, the average age in the Control group is 42.995 and the median is 43; the minimum age is 0 and the maximum is 102. The standard deviation is 20.985. When we divide the sample between Males and Females, we see that the average age for male patients is 43.465 and the median is 43. The minimum age is 0 and the maximum is 100. the standard deviation is 20.619. For female patients, the average and median ages are, respectively, 42.519 and 43; the minimum and maximum are 0 and 103 and the standard deviation is 21.313.
See Figures 1 2 3 4 5 for a complete description of the distributions.

Figure 1: Age Distribution with, potentially, two modes.

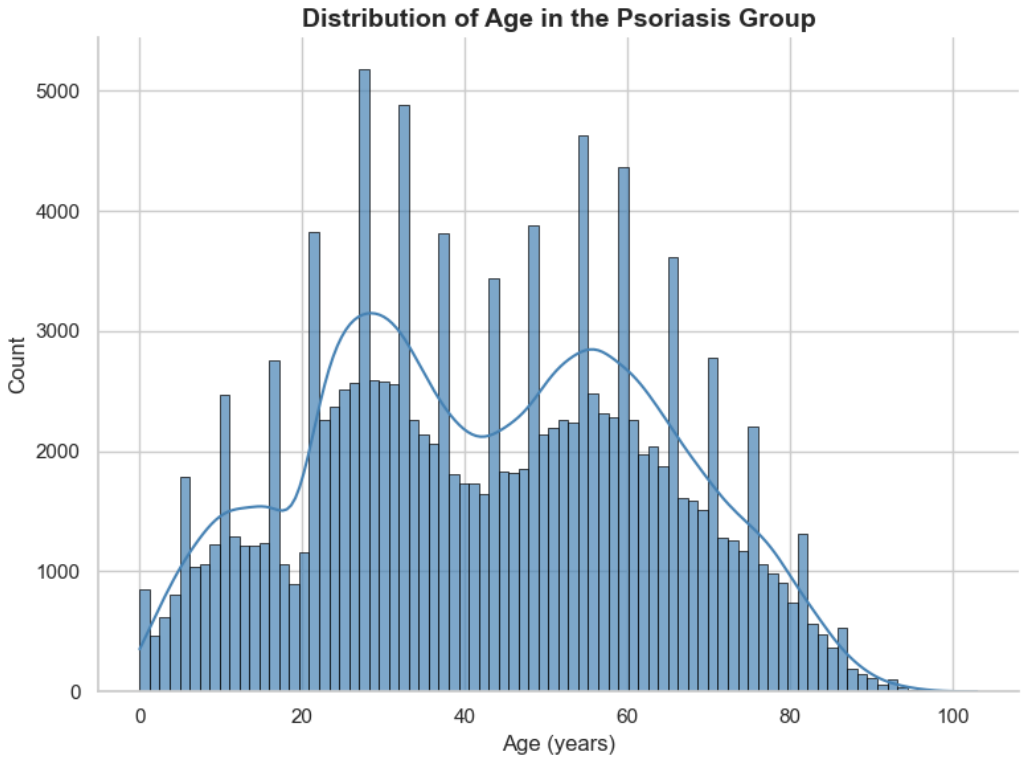

Figure 2: Age Distribution in Psoriatic patients.

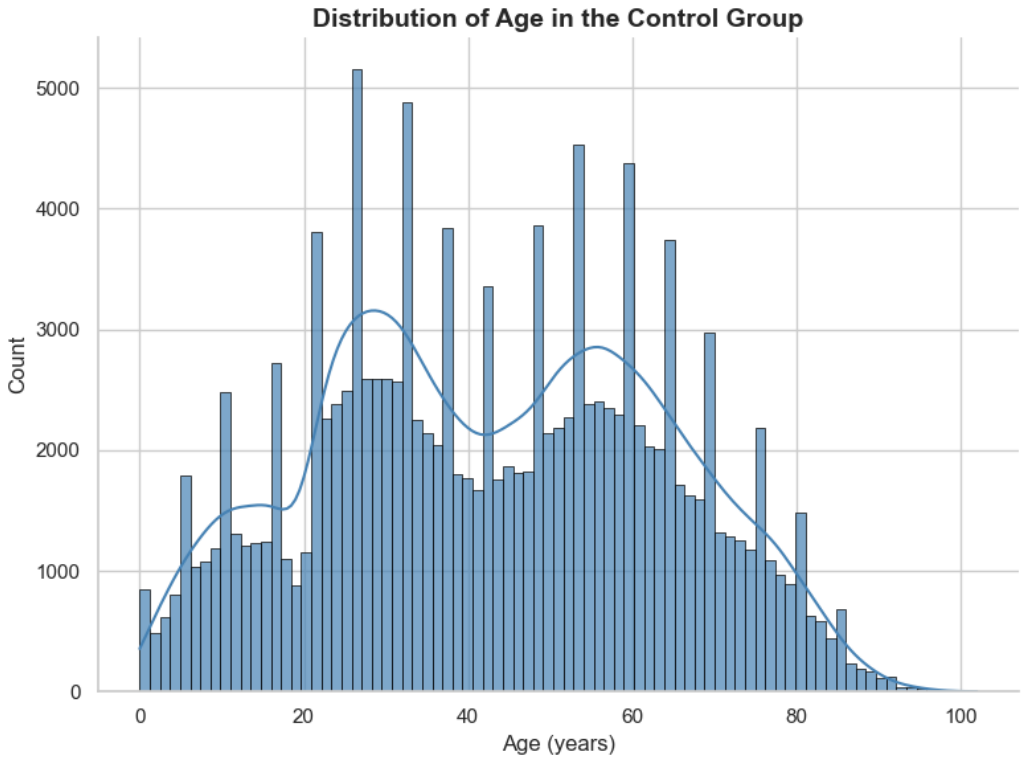

Figure 3: Age Distribution in Control patients.

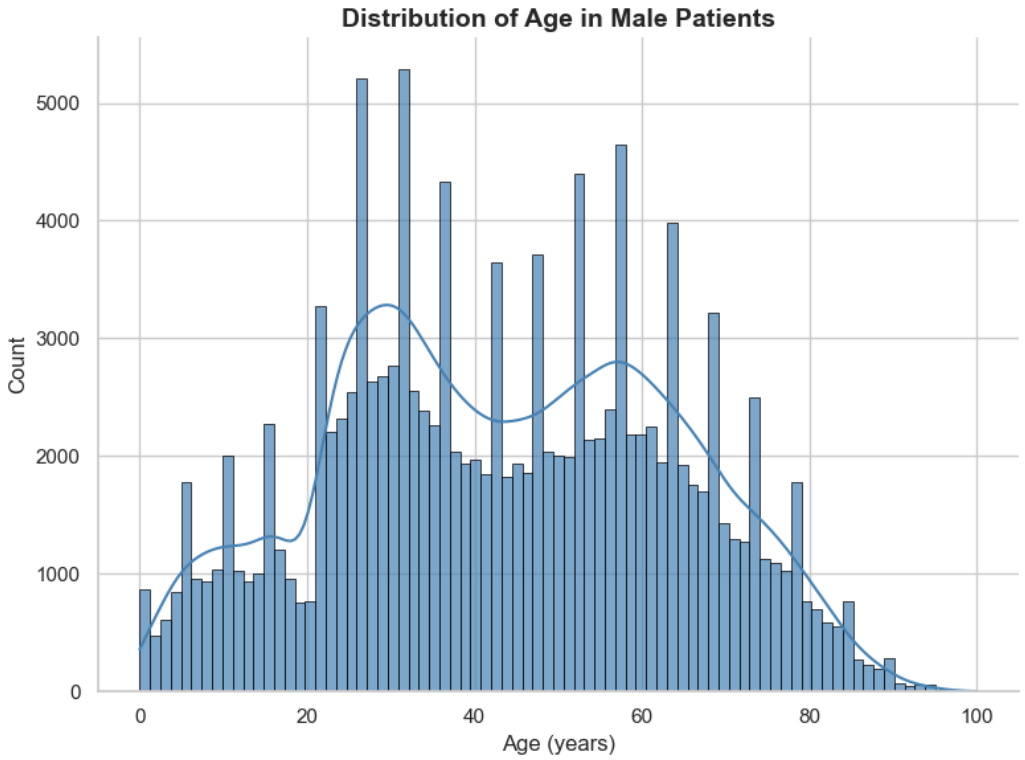

Figure 4: Age Distribution in Male patients.

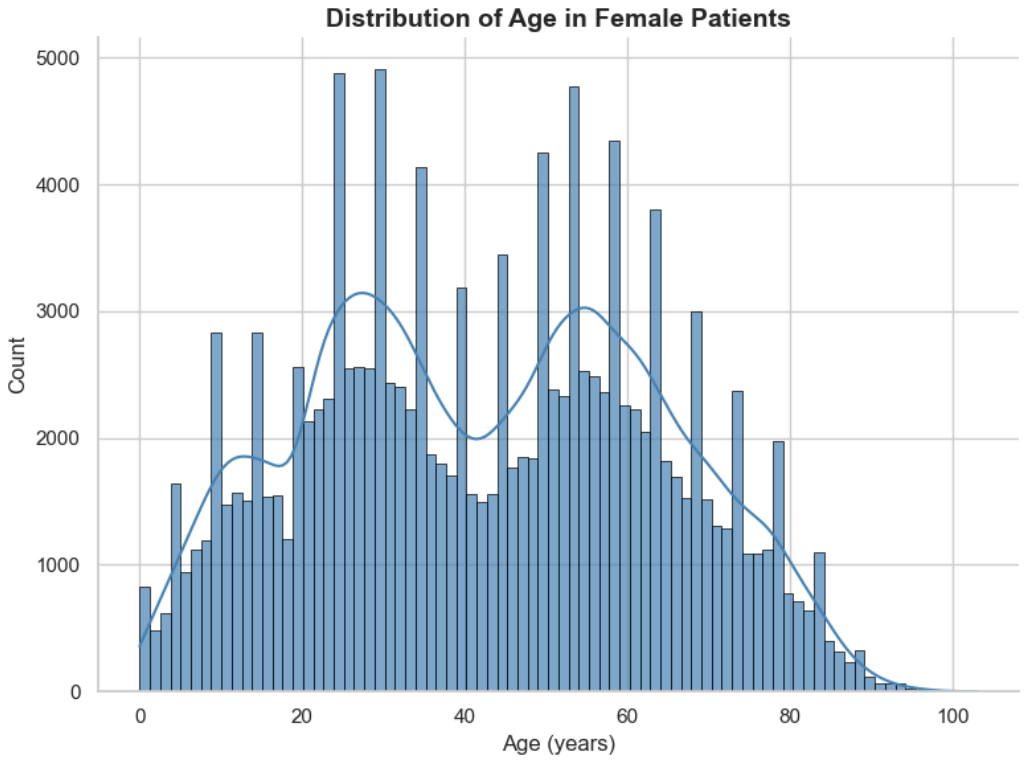

Figure 5: Age Distribution in Female patients.

Finally, we provide a quick summary of age stratification within male subjects divided between those with an Infertility diagnosis versus patients without (6 7).

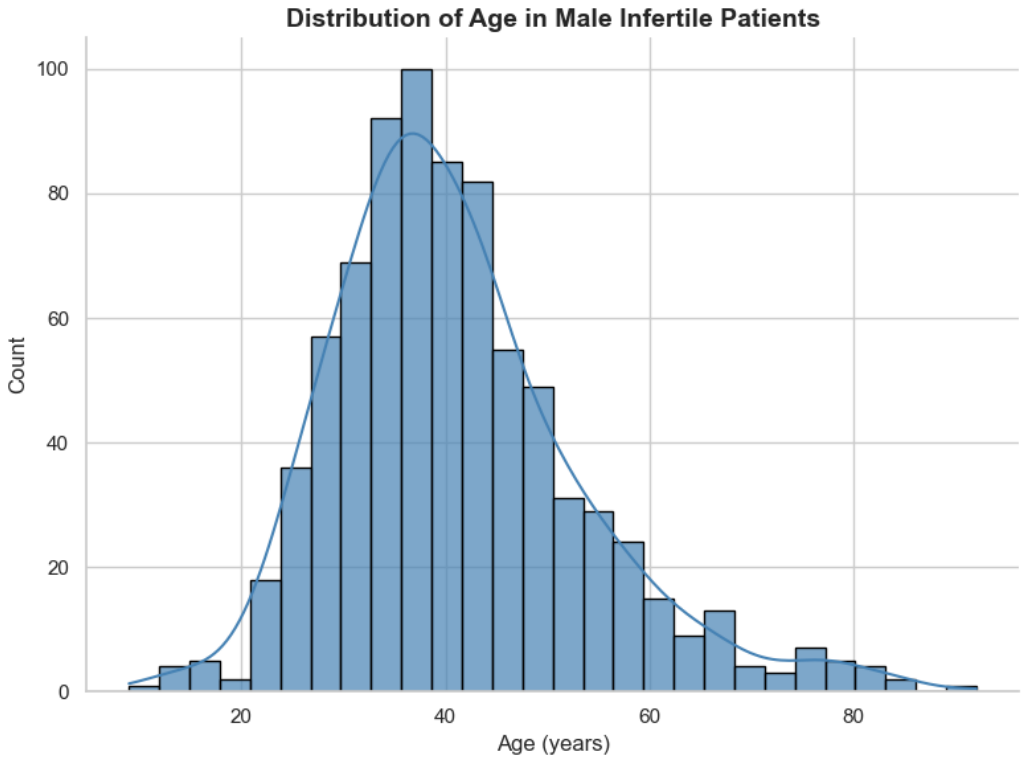

Figure 6: Age Distribution in Male Infertile patients.

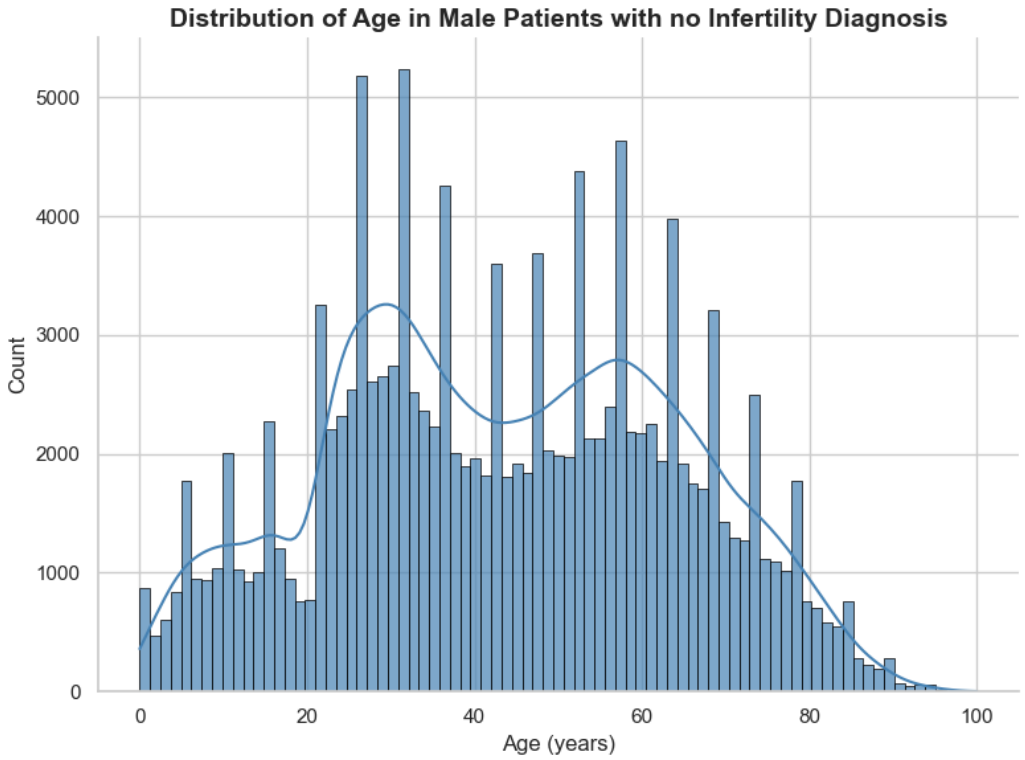

Figure 7: Age Distribution in Male patients without diagnosis.

Among male infertile subjects, the average and median age are, respectively, 40.95 and 39; the minimum and maximum ages are 9 and 92 and the standard deviation is 12.33. On the other hand, among patients without an infertility diagnosis, mean, median, minimum, maximum and standard deviation are 43.479, 43, 0, 100, 20.65. It seems that Infertility among male subjects is particularly concentrated around the mean and median in an age range that goes form the early thirties to the late fifties.
Lastly, if we take all the male patients with a Psoriasis diagnosis that is chronologically prior to the Infertility one, the average time span that intercurres between the two is of approximately 5 years.

3 Results

3.1 Comorbidity Analysis in Psoriasis vs. Control Groups

3.1.1 Statistical Methods and Robustness of Results

The tables 8 in Figure 8 present the prevalence of each comorbidity in Psoriasis and Control groups, along with 95% confidence intervals (CIs) computed using the Wilson score method. This method was selected to provide more accurate coverage and to mitigate biases that arise from the normal approximation, particularly when dealing with low-prevalence conditions. Indeed,the Wilson score method calculates 95% confidence intervals for a proportion by adjusting the standard error using the normal distribution. It provides more accurate bounds, especially for small samples or extreme proportions. The method is based on the following steps:

1. Given an observed proportion $\hat{p} = \frac{x}{n}$, where x is the number of successes and n is the sample size, the 95% confidence interval is centered around a "shrunk" estimate rather than the observed proportion.

2. The formula for the Wilson score interval is:

$$\hat{p} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}$$

3. The bounds of the interval are more precisely calculated as:

$$\frac{\hat{p} + \frac{z^2}{2n} \pm z\sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2}{4n^2}}}{1 + \frac{z^2}{n}}$$

4. Here, $z = 1.96$ for a 95% confidence level, $\hat{p}$ is the observed proportion, and n is the sample size.

5. This method avoids issues with small samples and extreme proportions, making it more reliable than the Wald interval.

In essence, the Wilson method improves accuracy by adjusting for the limitations of the normal approximation, ensuring more reliable and realistic confidence intervals across various sample sizes and proportions.

To assess statistical significance in the difference of prevalence rates between groups, we conducted chi-square tests (χ^2) on 2×2 contingency tables. Given that the minimum occurrence of each comorbidity was at least 100 cases in both groups, the assumptions for the χ^2 approximation hold, ensuring the robustness of the resulting p-values.
A significance threshold of $p < 0.05$ was applied to determine whether a given comorbidity showed a statistically significant difference between Psoriasis and Control patients. Only comorbidities meeting this threshold were retained in the tables. This combined approach ensures that the reported findings are statistically robust and clinically meaningful.

Comorbidity	Psoriasis Prevalence (%)	Psoriasis CI (%)	Control Prevalence (%)	Control CI (%)	Chi2	p-value	dof
Hyperlipidemia	46.085933	[45.83, 46.35]	42.386878	[42.13, 42.65]	388.757687	1.542878e-86	1
Smoking	39.061508	[38.81, 39.32]	35.277590	[35.03, 35.53]	429.692589	1.895815e-95	1
Hypertension	29.318595	[29.08, 29.56]	26.524514	[26.29, 26.76]	271.835825	4.516521e-61	1
Obesity	27.520057	[27.29, 27.75]	22.482439	[22.26, 22.70]	948.652692	2.603845e-208	1
Arthropathy	25.356392	[25.13, 25.58]	21.798538	[21.58, 22.02]	492.355620	4.377826e-109	1
Diabetes	18.441077	[18.24, 18.64]	16.275272	[16.08, 16.47]	229.110050	9.320245e-52	1
IHD	12.237475	[12.07, 12.41]	10.814049	[10.65, 10.98]	139.169437	4.044290e-32	1
Malignancy	11.172045	[11.01, 11.34]	10.151542	[9.99, 10.31]	76.550694	2.146320e-18	1
Reflux Esophagitis / Gastritis / Deudenitis	10.279194	[10.12, 10.44]	8.557675	[8.41, 8.71]	243.355210	7.296938e-55	1
Hypothyroidism	9.409877	[9.26, 9.56]	8.006418	[7.87, 8.15]	173.538354	1.248533e-39	1
Other Neurological Disease	8.691032	[8.54, 8.84]	7.900160	[7.76, 8.04]	57.542175	3.308094e-14	1
Depression	8.509895	[8.36, 8.66]	7.478695	[7.34, 7.62]	101.224427	8.213023e-24	1
other kidney Disease	8.283116	[8.14, 8.43]	7.520057	[7.38, 7.66]	55.992970	7.273061e-14	1
Asthma	8.250312	[8.11, 8.40]	7.129970	[7.00, 7.27]	123.811641	9.263255e-29	1
Arrhythmia	7.950793	[7.81, 8.09]	7.330362	[7.20, 7.47]	38.157218	6.526785e-10	1
Osteoporosis	7.817436	[7.68, 7.96]	7.445177	[7.31, 7.58]	13.730814	2.109647e-04	1
Anxiety	6.211446	[6.09, 6.34]	5.283651	[5.17, 5.40]	111.238713	5.245809e-26	1
Prostatic Hypertrophy	5.928329	[5.81, 6.05]	5.100374	[4.99, 5.22]	92.087399	8.293175e-22	1
Disability	5.855589	[5.73, 5.98]	6.076663	[5.95, 6.20]	6.068548	1.376101e-02	1
Chronic Renal Failure	5.646639	[5.53, 5.77]	5.413443	[5.30, 5.53]	7.253668	7.075633e-03	1
Peptic Ulcer	5.298627	[5.18, 5.42]	4.837226	[4.73, 4.95]	30.929031	2.676382e-08	1
COPD	4.600464	[4.49, 4.71]	3.679087	[3.58, 3.78]	149.755169	1.960985e-34	1
Glaucoma	4.319487	[4.21, 4.43]	4.002853	[3.90, 4.11]	17.546685	2.803390e-05	1
CHF	4.108397	[4.01, 4.21]	3.868069	[3.77, 3.97]	10.512793	1.185509e-03	1
Retinopathy	3.919415	[3.82, 4.02]	3.710465	[3.61, 3.81]	8.285427	3.996451e-03	1
PVD	3.214833	[3.12, 3.31]	2.702799	[2.62, 2.79]	63.842340	1.347854e-15	1
Dementia / Alzheimers / OMS	2.831877	[2.75, 2.92]	3.314673	[3.22, 3.41]	54.700711	1.403555e-13	1
Deafness	2.725620	[2.64, 2.81]	2.442503	[2.36, 2.52]	22.212738	2.440493e-06	1
Neuroses	2.632911	[2.55, 2.72]	2.353361	[2.28, 2.43]	22.424197	2.186025e-06	1
Carotid Artery Disease	2.208593	[2.13, 2.29]	1.822785	[1.75, 1.89]	52.644377	3.997538e-13	1
Other Endocrine and Metabolic Disease	2.194331	[2.12, 2.27]	1.829916	[1.76, 1.90]	47.038986	6.958860e-12	1
OncBreat	2.145124	[2.07, 2.22]	1.991799	[1.92, 2.07]	8.061245	4.522211e-03	1
Joint Replacement	2.142271	[2.07, 2.22]	1.842040	[1.77, 1.91]	32.215069	1.380161e-08	1
Other Hematologic Dis (excl. Iron Def Anemia)	1.975397	[1.90, 2.05]	1.799964	[1.73, 1.87]	11.556456	6.751440e-04	1
Gout	1.513995	[1.45, 1.58]	1.241576	[1.18, 1.30]	38.092378	6.747320e-10	1
Rheumatoid Arthritis	1.399893	[1.34, 1.46]	0.801569	[0.76, 0.85]	230.016148	5.913179e-52	1
Hyperthyroidism	1.345694	[1.29, 1.41]	1.256552	[1.20, 1.32]	4.269306	3.880677e-02	1

Comorbidity	Psoriasis Prevalence (%)	Psoriasis CI (%)	Control Prevalence (%)	Control CI (%)	Chi2	p-value	dof
Irritable Bowel Syndrome	1.277946	[1.22, 1.34]	0.920663	[0.87, 0.97]	81.991180	1.366953e-19	1
OncUnKnow	1.274380	[1.22, 1.33]	0.932073	[0.88, 0.98]	74.983999	4.745445e-18	1
Parkinsons Disease	1.053307	[1.00, 1.11]	1.187377	[1.13, 1.25]	11.255637	7.938158e-04	1
Other Liver Disease	0.967730	[0.92, 1.02]	0.636834	[0.60, 0.68]	96.045668	1.122638e-22	1
Chronic Bronchitis	0.939205	[0.89, 0.99]	0.713853	[0.67, 0.76]	43.163054	5.036277e-11	1
Congenital Anomalies	0.935639	[0.89, 0.99]	0.847923	[0.80, 0.90]	6.004769	1.426726e-02	1
OnclymphomaNonHodgkin	0.887859	[0.84, 0.94]	0.675343	[0.63, 0.72]	40.558336	1.908326e-10	1
OncUrineBlader	0.872883	[0.83, 0.92]	0.699590	[0.66, 0.74]	26.770114	2.291541e-07	1
Cardiomyopathy	0.831521	[0.79, 0.88]	0.728829	[0.69, 0.77]	9.419466	2.146940e-03	1
Other Rheumatic / Autoimmune	0.815832	[0.77, 0.86]	0.606169	[0.57, 0.65]	43.361965	4.549459e-11	1
Chronic Act/Per Hepatitis	0.769478	[0.73, 0.82]	0.552683	[0.52, 0.59]	49.848959	1.660484e-12	1
Aortic Aneurism	0.748083	[0.70, 0.79]	0.615439	[0.58, 0.66]	18.022978	2.182546e-05	1
OncMelanoma	0.742378	[0.70, 0.79]	0.653236	[0.61, 0.70]	7.912135	4.910428e-03	1
Crohns Disease	0.717418	[0.67, 0.76]	0.400784	[0.37, 0.44]	125.862500	3.295435e-29	1
Hepatitis B Carrier	0.683188	[0.64, 0.73]	0.760920	[0.72, 0.81]	5.801893	1.600893e-02	1
Blindness	0.536994	[0.50, 0.58]	0.667499	[0.63, 0.71]	19.730430	8.917011e-06	1
Celiac Disease	0.535568	[0.50, 0.58]	0.366554	[0.34, 0.40]	44.227953	2.922776e-11	1
Ulcerative Colitis	0.523444	[0.49, 0.56]	0.375825	[0.35, 0.41]	33.804654	6.093292e-09	1
Bronchiectasis	0.518452	[0.48, 0.56]	0.446425	[0.41, 0.48]	7.426813	6.425905e-03	1
Dialysis	0.500624	[0.47, 0.54]	0.592619	[0.55, 0.63]	10.746282	1.044890e-03	1
Cirrhosis	0.472098	[0.44, 0.51]	0.368693	[0.34, 0.40]	17.662037	2.638403e-05	1
OncThyroid	0.446425	[0.41, 0.48]	0.375825	[0.35, 0.41]	8.363961	3.827353e-03	1
OncKidney	0.432163	[0.40, 0.47]	0.371546	[0.34, 0.40]	6.286131	1.216864e-02	1
Amputation of Limb	0.417900	[0.39, 0.45]	0.491353	[0.46, 0.53]	8.197267	4.195354e-03	1
Polymyalgia Rheumatica	0.392227	[0.36, 0.43]	0.295953	[0.27, 0.33]	18.671501	1.552862e-05	1
Hidradenitis Suppurativa	0.281690	[0.26, 0.31]	0.151186	[0.13, 0.17]	54.688383	1.412387e-13	1
SLE	0.278124	[0.25, 0.31]	0.224639	[0.20, 0.25]	7.786951	5.262493e-03	1
OncSaracoma	0.235336	[0.21, 0.26]	0.195400	[0.17, 0.22]	5.019088	2.506937e-02	1
Sarcoidosis	0.234623	[0.21, 0.26]	0.157604	[0.14, 0.18]	20.857268	4.948112e-06	1
OncGastro	0.206097	[0.18, 0.23]	0.242467	[0.22, 0.27]	3.983497	4.594807e-02	1
OncGenitalia	0.147620	[0.13, 0.17]	0.100553	[0.09, 0.12]	12.155888	4.893313e-04	1
Cerebral Palsy	0.128365	[0.11, 0.15]	0.165448	[0.15, 0.19]	6.322395	1.192225e-02	1
Pemphigus Vulgaris	0.121947	[0.10, 0.14]	0.072740	[0.06, 0.09]	16.954233	3.829185e-05	1
Myasthenia Gravis	0.106971	[0.09, 0.13]	0.077019	[0.06, 0.09]	6.521503	1.065779e-02	1
Kidney Transplant	0.091282	[0.08, 0.11]	0.163309	[0.14, 0.19]	28.046907	1.184102e-07	1

Figure 8: Table of differences in Comorbidity Prevalence, Psoriasis versus Control. Statistically significant intervals at 5% alpha error.

The analysis reveals that psoriasis is strongly associated with a range of systemic comorbidities, reinforcing its classification as a multisystem inflammatory disease rather than an isolated dermatological condition. Below, we summarize key findings across major disease categories.

Metabolic and Cardiovascular Comorbidities Psoriasis patients exhibited significantly higher prevalence of metabolic syndrome components:

- **Hyperlipidemia (46.1% vs. 42.4%,** $p < 10^{-86}$**)**, **Obesity (27.5% vs. 22.5%,** $p < 10^{-208}$**)**, and **Diabetes (18.4% vs. 16.3%,** $p < 10^{-22}$**)**.
- Increased **hypertension (29.3% vs. 26.5%,** $p < 10^{-61}$**)** and **ischemic heart disease (12.2% vs. 10.8%,** $p < 10^{-22}$**)** confirm the pro-atherogenic state linked to psoriasis.
- **Chronic renal failure (5.7% vs. 5.1%,** $p = 0.0073$**)** and **peripheral vascular disease (PVD, 3.2% vs. 2.7%,** $p < 10^{-15}$**)** highlight potential vascular complications.

Autoimmune and Rheumatic Diseases

- **Arthropathy (25.4% vs. 21.8%,** $p < 10^{-109}$**)** had one of the largest relative differences, confirming the strong psoriasis-arthritis link.
- Increased prevalence of Crohn's Disease (0.72% vs. 0.40%, $p < 10^{-29}$) and Ulcerative Colitis (0.52% vs. 0.38%, $p < 10^{-9}$) aligns with the inflammatory overlap between psoriasis and inflammatory bowel disease (IBD).

Neurological and Psychiatric Conditions

- **Depression (8.5% vs. 7.5%,** $p < 10^{-24}$**)** and **Anxiety (6.2% vs. 5.3%,** $p < 10^{-13}$**)** emphasize the psychological burden of psoriasis.
- Higher prevalence of neurological disorders (8.7% vs. 7.9%, $p < 10^{-14}$) suggests an increased risk of neuroinflammation and cognitive decline.

Pulmonary and Respiratory Disorders

- **COPD (4.6% vs. 3.7%,** $p < 10^{-34}$**)** and **Asthma (8.3% vs. 7.1%,** $p < 10^{-29}$**)** suggest increased chronic respiratory inflammation.
- Higher smoking rates among psoriasis patients could partially explain these associations.

Gastrointestinal and Hepatic Conditions

- **Reflux Esophagitis/Gastritis (10.3% vs. 8.6%,** $p < 10^{-52}$**)** and **Peptic Ulcer Disease (5.3% vs. 4.8%,** $p < 10^{-8}$**)** were significantly more common in psoriasis patients.
- Higher prevalence of cirrhosis (0.47% vs. 0.37%, $p < 10^{-5}$) suggests possible systemic inflammation and hepatotoxic effects from psoriasis treatments.

Oncological Risks

- Overall malignancy prevalence (11.2% vs. 10.2%, $p < 10^{-18}$) was significantly higher in psoriasis patients.
- Specific cancers with notable differences:
  - **Non-Hodgkin's Lymphoma (0.88% vs. 0.68%,** $p < 10^{-7}$**)**
  - **OncKidney (0.43% vs. 0.38%,** $p = 0.0122$**)**
  - **OncBreast (2.15% vs. 1.99%,** $p = 0.0045$**)**

The findings reinforce that psoriasis is not only a dermatological condition but a systemic inflammatory disorder with elevated risk for metabolic, cardiovascular, autoimmune, psychiatric, pulmonary, gastrointestinal, and oncologic conditions.

- The strongest associations were observed for cardiovascular/metabolic diseases, psoriatic arthritis, and mental health conditions, emphasizing the need for multidisciplinary care.
- The increased prevalence of depression and anxiety highlights the need for psychological support in psoriatic patients.
- Higher rates of neurological and oncological conditions warrant further investigation into neuroinflammation and immune dysregulation.

Overall, these findings support the importance of systematic comorbidity screening and early intervention in psoriasis patients to improve long-term outcomes. Clinicians should adopt a comprehensive, multidisciplinary approach that addresses both dermatological and systemic health risks.

3.2 Longitudinal Analysis: Comorbidities in Patients Developing Psoriasis First

3.2.1 Methodology and Statistical Considerations

We go further into the details of the study.This analysis examines comorbidities that developed after psoriasis diagnosis to assess whether psoriasis increases the risk of specific conditions. A matched control group without psoriasis was included for comparison.

- **Statistical Testing:** All comparisons were performed using the **Chi-Square Test**.
- **Multiple Testing Correction:** Only statistically significant results ($p < 0.05$) were retained.
- **Confidence Intervals:** 95% Wilson confidence intervals were used for prevalence estimates.

Metabolic and Cardiovascular Disorders Hyperlipidemia (**15.69%** vs. **0.23%**, $p < 0.00001$), Obesity (**11.30%** vs. **0.21%**, $p < 0.00001$), and Hypertension (**6.09%** vs. **0.14%**, $p < 0.00001$) were among the most prevalent comorbidities in psoriasis patients, aligning with prior findings linking psoriasis to systemic inflammation and cardiovascular risk.

Autoimmune and Immune-Mediated Diseases Hypothyroidism was significantly more common in psoriasis patients (**3.10%** vs. **0.05%**, $p < 0.00001$), suggesting shared autoimmune mechanisms. Similarly, Rheumatoid Arthritis (**0.36%** vs. **0.02%**, $p < 0.00001$) supports the concept of a psoriatic disease spectrum.

Neuropsychiatric and Mental Health Conditions Depression (**2.11%** vs. **0.03%**, $p < 0.00001$) and Anxiety (**1.85%** vs. **0.04%**, $p < 0.00001$) showed a **70-fold and 46-fold increased prevalence**, respectively, reinforcing the need for integrated psychiatric care in psoriasis management.

Oncological Risks Malignancy prevalence (**2.72%** vs. **0.06%**, $p < 0.00001$) was significantly higher in psoriasis patients, supporting the hypothesis that chronic inflammation may contribute to oncogenesis.

These findings confirm that psoriasis is not only associated with systemic diseases but may actively contribute to their development over time.

3.3 Pre-Existing Comorbidities as Predictors of Psoriasis Onset

3.3.1 Introduction and Methodological Framework

This analysis aims to determine whether specific pre-existing comorbidities increase the likelihood of later developing psoriasis. Unlike previous sections, which explored psoriasis as a precursor to other conditions, this study investigates whether certain conditions act as risk factors that predispose individuals to psoriasis onset.
A matched control group of individuals without psoriasis was included for comparison. By examining the prevalence of various comorbidities in individuals who subsequently developed psoriasis, we can identify potential causal relationships.

To ensure statistical robustness, the following methodology was applied:

- **Statistical Test Used: Chi-Square Test** was applied to assess differences in comorbidity prevalence between psoriasis and control groups.
- **Significance Threshold:** Only results with $p < 0.05$ were retained, ensuring that differences are statistically meaningful.
- **Confidence Intervals: Wilson 95% confidence intervals** were used for prevalence estimates to account for low event frequencies in certain comorbidities.

This analysis is crucial for understanding the risk profile of psoriasis and identifying modifiable factors that could be targeted for prevention.

Smoking: The Strongest Predictive Factor for Psoriasis Development

- **Smoking Prevalence Before Psoriasis:** 30.63% [CI: 30.29%, 30.98%]
- **Smoking Prevalence in Controls:** 0.12% [CI: 0.07%, 0.15%]
- **Chi-Square Value:** 11683.96 ($p < 0.00001$)

Smoking emerges as the **single most significant predictor** of future psoriasis development. Individuals who later developed psoriasis were **255 times more likely** to have been smokers compared to controls.
Nicotine is known to influence **keratinocyte proliferation, immune responses, and vascular function**, all of which are implicated in psoriasis pathogenesis. This finding strongly supports **smoking cessation as a preventive measure** for high-risk individuals.

Metabolic Dysregulation and Psoriasis Onset

- **Hyperlipidemia:** 24.70% (Psoriasis) vs. 0.10% (Control), $p < 0.00001$
- **Obesity:** 15.50% (Psoriasis) vs. 0.05% (Control), $p < 0.00001$
- **Hypertension:** 12.53% (Psoriasis) vs. 0.06% (Control), $p < 0.00001$

All three conditions show massive risk multipliers for psoriasis onset:

- **Hyperlipidemia:** 247-fold higher risk
- **Obesity:** 310-fold increased risk
- **Hypertension:** 209-fold increased risk

These findings reinforce the concept of **psoriasis as a systemic inflammatory disorder** rather than just a skin disease. Chronic low-grade inflammation driven by **dyslipidemia, insulin resistance, and vascular dysfunction** may serve as a trigger for immune dysregulation, ultimately leading to psoriasis onset.

Autoimmune and Inflammatory Conditions Preceding Psoriasis

- **Hypothyroidism:** 4.70% (Psoriasis) vs. 0.03% (Control), $p < 0.00001$
- **Rheumatoid Arthritis:** 0.48% (Psoriasis) vs. 0.00% (Control), $p < 0.00001$

The strong association between **hypothyroidism and rheumatoid arthritis with later psoriasis onset** suggests shared **autoimmune pathophysiology**.
Thyroid dysfunction may influence immune responses, leading to systemic inflammation. Rheumatoid arthritis and psoriasis share overlapping genetic markers (e.g., **HLA-Cw6**) and inflammatory pathways.

Mental Health and Neurological Disorders as Predictors

- **Depression:** 3.30% (Psoriasis) vs. 0.02% (Control), $p < 0.00001$
- **Anxiety:** 2.72% (Psoriasis) vs. 0.01% (Control), $p < 0.00001$

Individuals with a history of **depression were 165 times more likely** to develop psoriasis, while **anxiety disorders were associated with a 272-fold increased risk**.
Psychological stress is a well-known psoriasis trigger. Dysregulation of the **hypothalamic-pituitary-adrenal (HPA) axis** and **pro-inflammatory cytokine production** may explain this association. These findings suggest that **early psychiatric intervention may help reduce psoriasis risk in susceptible individuals**.

3.3.2 Interpretation in Relation to Previous Analyses

Key Takeaways:

- Some conditions (e.g., smoking, metabolic syndrome, depression) act as both risk factors and consequences of psoriasis.
- Other conditions (e.g., rheumatoid arthritis, thyroid disorders) appear to be more predisposing than resulting from psoriasis.
- Psoriasis should be considered part of a larger systemic disease framework rather than an isolated dermatological disorder.

3.4 Analysis of Psoriasis and Infertility: Data Processing, Clustering, and Classification

3.4.1 Introduction

This section presents the analysis on the dataset investigating the relationship between psoriasis and infertility. The study includes extensive preprocessing, statistical tests, clustering, and classification methods to uncover potential associations and subgroup structures within the data.

3.4.2 Data Import and Preprocessing

The dataset was imported and preprocessed using Python libraries such as `numpy`, `pandas`, `matplotlib`, and `seaborn`. Several additional libraries, including `scikit-learn` and `scipy.stats`, were used for clustering and statistical analysis.
Key preprocessing steps included:

- Removing missing values and filtering for male patients.
- Binarizing categorical medical conditions.
- Creating time-based features such as the difference in days and months between psoriasis and infertility diagnoses.

A function was implemented to compute these time differences:

```
def days_between(d1, d2):
    if pd.isna(d1) or pd.isna(d2):
        return np.nan
    else:
        return (d2 - d1).days
```

3.4.3 Chi-Squared Test for Feature Importance

A chi-squared test was conducted to identify statistically significant comorbidities associated with infertility in psoriasis patients. The function `run_chi2` was applied, comparing two groups: patients with only psoriasis and those with both psoriasis and infertility.
The results identified several comorbidities with p-values below 0.05, indicating significant associations. Key findings include:

- Hyperthyroidism (p = 1.831e-11)
- Diabetes (p = 4.067e-02)
- Depression (p = 1.295e-05)
- Multiple Sclerosis (p = 2.353e-80)
- Joint Replacement (p = 6.422e-44)

3.4.4 Clustering Analysis

Several clustering techniques were applied to identify distinct subgroups within the dataset. Clustering methods included:

- **K-Means Clustering:** Used to explore potential groupings based on silhouette scores.
- **Agglomerative Clustering:** A hierarchical method for grouping similar patients.
- **Gaussian Mixture Model (GMM):** Applied to capture probabilistic cluster memberships.
- **Spectral Clustering:** Used to analyze network-based relationships between features.

The silhouette scores for different numbers of clusters suggested that the optimal number of clusters was **two**, confirming the presence of two distinct subgroups.

3.4.5 Dimensionality Reduction and Multiple Correspondence Analysis (MCA)

To reduce the complexity of the data while preserving important categorical relationships, Multiple Correspondence Analysis (MCA) was applied. Silhouette scores were computed for different numbers of components (`n_components` = 10, 20, ..., 70), with the best results obtained for 10 components.
The MCA transformation facilitated a more structured clustering approach, which yielded, on average, silhouettes scores in the order of 0.8; these were then used for downstream classification tasks.

3.4.6 Classification Models

Several classification models were trained to predict infertility based on psoriasis-related features. These included:

- **Random Forest Classifier:** Trained on the transformed dataset, achieving an AUC-ROC score of 0.6088.
- **StepMix Model:** A latent class model used for categorical feature learning, which achieved a test accuracy of 0.4817.
- **Spectral Clustering-based Classification:** Used to improve subgroup identification, but with mixed classification performance.

3.4.7 Handling Imbalanced Data

Given the highly imbalanced nature of the dataset (with infertility cases being much rarer than psoriasis cases), **Synthetic Minority Over-sampling Technique (SMOTE)** was applied to balance the dataset before training machine learning models. Still,no substantial improvement was obtained in AUC-ROC terms.

3.4.8 Final Results and Conclusion

The study confirmed that:

- Patients with psoriasis and infertility form a distinct subgroup with unique comorbidity profiles.
- Clustering approaches consistently identified two major patient subgroups.
- Significant medical conditions, such as hyperthyroidism and multiple sclerosis, show strong associations with psoriasis-related infertility.
- Classification models provided moderate predictive accuracy, with further improvements needed.

3.5 Analysis of Infertility and Psoriasis: Chronological inversion

3.5.1 Chi-Squared Test for Feature Importance

Chi-squared tests were performed to identify significant comorbidities associated with infertility in psoriasis patients. The function `run_chi2` was applied, comparing two groups: patients with infertility alone and those with both infertility and psoriasis.
Statistically significant conditions ($p < 0.05$) include:

- **Chronic Renal Failure** ($p = 0.0193$)
- **Arthropathy** ($p = 0.0072$)
- **Hyperthyroidism** ($p = 0.0058$)
- **Diabetes** ($p = 0.0246$)
- **COPD** ($p = 0.0256$)

3.5.2 Feature Engineering

The dataset was transformed by assigning categorical labels to comorbidities based on their relative timing with respect to infertility and psoriasis diagnoses:

- **-1**: Diagnosed before infertility
- **0**: Not diagnosed or diagnosed after psoriasis

- **1**: Diagnosed after infertility but before psoriasis

Additionally, time differences were computed and binned into:

- ≤ -24 months
- -24 to -12 months
- -12 to 0 months
- 0 to 12 months
- 12 to 24 months
- > 24 months

3.5.3 Dimensionality Reduction and Clustering

Multiple Correspondence Analysis (MCA) was applied to reduce the dataset's dimensionality while preserving categorical relationships. Silhouette scores were computed for different numbers of MCA components.
Optimal MCA Components and Clustering Scores:

- **10 components:** Best silhouette score ~ 0.85
- **20 components:** Stable performance (~ 0.83)
- **30 components:** Slight drop in silhouette score (~ 0.79)

Several clustering algorithms were evaluated:

- **K-Means:** Identified two distinct clusters.
- **Agglomerative Clustering:** Produced similar subgroup separation.
- **Gaussian Mixture Model (GMM):** Confirmed two primary subgroups.
- **Spectral Clustering:** Achieved moderate performance.

3.5.4 Subgroup Analysis

Patients were clustered into two primary groups based on MCA-transformed features. Statistical tests identified key differences:

- **Group 1** (Infertility-only) vs. **Group 2** (Infertility + Psoriasis)
  - Significant features: **Hepatitis C Carrier**, **Malignancy**, **Hypothyroidism**, **Diabetes**, **Cardiomyopathy**, **Arrhythmia**.
- Within-group comparisons:
  - Patients who developed psoriasis after infertility showed distinct medical profiles.
  - Specific comorbidities had different time associations with psoriasis onset.

3.5.5 StepMix Model and Classification

A StepMix latent class model was applied to categorize patients into four latent classes. The model fit results:

- **4 latent classes identified** (Entropy = 0.51)
- **Best classification accuracy**: 0.4817
- **AIC = 444.91**, **BIC = 810.21**

Random Forest classification was also tested, yielding a moderate AUC-ROC score of 0.6088.

3.5.6 Final Results and Conclusion

The analysis confirmed:

- The existence of two primary patient subgroups.
- Comorbidities such as **chronic renal failure**, **arthropathy**, and **cardiovascular diseases** are strongly associated with infertility and psoriasis.
- StepMix provided a probabilistic clustering approach but requires further refinement.

3.6 Comorbidome Analysis: Psoriasis and Male Infertility

3.6.1 Introduction

This section presents an analysis aimed at establishing a **comorbidome** for male patients transitioning from **psoriasis to potential infertility**. The objective is to compute **odds ratios (ORs)** comparing male patients with and without infertility, identifying significant comorbidities associated with infertility in the context of psoriasis.
The study employed **logistic regression models** to estimate ORs and their statistical significance (p-values). Several cohort definitions were used:

1. Patients with psoriasis, comparing **fertile vs. infertile males** (infertility diagnosed after psoriasis).

2. Patients with infertility, comparing **psoriasis vs. control** (all males had infertility diagnosed after psoriasis or after a control reference date).

3.6.2 Data Processing

The dataset underwent several preprocessing steps to filter and prepare it for the comorbidome analysis:

- **Filtering for male patients** to ensure comparability.
- **Identifying psoriasis and infertility diagnoses** using clinical records.
- **Handling missing psoriasis diagnosis dates:**
  - If a patient was classified as **psoriatic** but lacked a recorded diagnosis date, the **'Start_D'** column was used as a proxy for the first psoriasis-related medical entry.
- **Filtering for comorbidities** recorded prior to infertility diagnosis, ensuring that conditions analyzed as potential risk factors preceded infertility.
- **Binary encoding of comorbidities** for logistic regression.

3.6.3 Methodology

For each comorbidity, a **logistic regression model** was fit using the following model:

$$\log\left(\frac{P(\text{Infertility}=1)}{P(\text{Infertility}=0)}\right) = \beta_0 + \beta_1 X_{\text{Comorbidity}} \tag{1}$$

where:

- $X_{\text{Comorbidity}}$ is a binary variable indicating the presence of a given comorbidity.
- β_1 represents the **log-odds ratio** of infertility for patients with the comorbidity compared to those without it.

For each regression:

- The **odds ratio (OR)** was computed as e^{β_1}.
- A **95% confidence interval (CI)** for the OR was estimated.
- A **p-value** was extracted to assess statistical significance.

3.6.4 Results

Psoriatic Patients: Fertile vs. Infertile In the first logistic regression analysis comparing **fertile vs. infertile males** among psoriatic patients, the following significant results were observed:

- **Hyperlipidemia**: OR < 1, p-value < 0.001
- **Hypertension**: OR < 1, p-value < 0.001

These results suggest that **hyperlipidemia and hypertension were less frequent among infertile males compared to fertile ones** in the psoriatic cohort. This unexpected inverse association might be influenced by confounders such as age, treatment history, or health-seeking behaviors.

3.6.5 Infertile Patients: Psoriasis vs. Control

In the second analysis comparing **psoriatic vs. control patients among those diagnosed with infertility**, no significant comorbidities were found:

- **All p-values were above 0.05, indicating no statistically significant ORs**.

This suggests that within the infertile male population, having a prior psoriasis diagnosis did not appear to be significantly associated with any specific comorbidity.

3.6.6 Psoriasis vs. Control (Ignoring Infertility)

To validate our approach, an additional analysis was conducted comparing **psoriatic vs. non-psoriatic males** without considering infertility. In this case, several significant ORs were obtained.
However, this analysis does not fully align with the study's objective, which is to establish a comorbidome in the context of infertility. Therefore, while it confirms that psoriasis is associated with various comorbidities, it does not directly contribute to the infertility-specific findings.

3.6.7 Challenges and Limitations

- **Dataset Imbalance:** The dataset contains a **highly imbalanced distribution**, with only 3% of patients diagnosed with infertility. This leads to statistical power issues and potential overfitting in logistic models.
- **Control Group Uncertainty:** Patients without an infertility diagnosis may not be true controls, as their fertility status is unknown rather than explicitly documented as "fertile."
- **Comorbidities with Low Prevalence:** Several conditions had very low frequencies, making statistical inference unreliable for those cases.

3.6.8 Conclusion

This analysis attempted to define a **comorbidome** for psoriasis-related infertility by computing **odds ratios and statistical significance** for various medical conditions. The key findings are:

- **Hypertension and hyperlipidemia were inversely associated with infertility** among psoriatic males.
- **No significant comorbidities emerged when comparing infertile psoriatic and control patients.**
- **The dataset's imbalance and control group definition present challenges that should be addressed in future studies.**

Future work should consider:

- Adjusting for confounders such as **age, BMI, and medication use**.
- Exploring alternative statistical methods like **propensity score matching** to create a more balanced cohort.
- Expanding the dataset to include **more documented infertility cases** for improved statistical power.

The findings provide valuable insights but require further refinement to robustly characterize the relationship between psoriasis, infertility, and associated comorbidities.

3.7 Refined Comorbidome Analysis: Excluding Psoriatic Patients Without a Diagnosis Date

3.7.1 Revised Logistic Regression Analysis

To validate the robustness of our initial results, we refined the analysis by **excluding psoriatic patients without a recorded diagnosis date**. This adjustment ensures that only confirmed psoriatic patients are included, minimizing biases introduced by uncertain disease onset.
After applying this filter, the sample size was reduced to **14,528 male psoriatic patients**. Logistic regression was repeated to compute **odds ratios (ORs)** for infertility, and the following significant comorbidities emerged ($p\text{-}value < 0.001$), as it is shown in :

- **Hyperprolactinemia**: OR = 33.59
- **Peripheral Vascular Disease (PVD)**: OR = 410.19

These odds ratios are markedly high, suggesting a strong association between these conditions and infertility among psoriatic patients. Hyperprolactinemia, a condition linked to hormonal imbalances affecting reproductive function, appears to be a potential infertility driver. Similarly, PVD, indicative of circulatory issues, may contribute to infertility through vascular dysfunction. However, the 95% confidence intervals also happened to be quiet wide, touching a minimum of 4.23 for Hyperprolactinemia. PVD on the other hand would be included in an interval going from 281 to 597. We suspect that the rareness of such diseases would inevutably impact the dataset, thus the statistical tests and scores.

Limitations in Control Group Selection A critical observation from this analysis is that **patients without an infertility diagnosis do not necessarily constitute a valid control group**. The dataset's design primarily focuses on psoriasis, meaning that many younger patients (e.g., those born in 2009 and 2013) have likely not yet reached an age where infertility could be assessed. This highlights an intrinsic limitation: **it is challenging to extract a definitive "non-infertile" control group** from this dataset.

3.7.2 Temporal Relationship Between Psoriasis and Prostate Conditions

To further investigate comorbid conditions, we examined **prostate-related diagnoses** within the psoriatic cohort. Specifically, we assessed whether **psoriasis preceded or followed the diagnosis of prostate cancer and benign prostatic hyperplasia (BPH).**

Prostate Cancer Among the **14,528 male psoriatic patients** with a recorded diagnosis:

- 14,168 patients did **not** have a prostate cancer diagnosis.
- 121 patients had a prostate cancer diagnosis **before** psoriasis.
- 239 patients had a prostate cancer diagnosis **after** psoriasis.

Benign Prostatic Hyperplasia (BPH) For **BPH**, the distribution was as follows:

- 1,054 patients had a BPH diagnosis **before** psoriasis.
- 894 patients had a BPH diagnosis **after** psoriasis.
- The remaining patients had no BPH diagnosis.

These findings suggest a **bidirectional relationship** between psoriasis and prostate conditions, warranting further investigation into whether psoriasis acts as a precursor to prostate disease or if pre-existing prostate conditions increase susceptibility to psoriasis.

3.7.3 Checking for Robustness

p-value threshold In both of the two previous analyses, we decided to lower the p-value threshold to 0.05 and check whether the profile of statistically significant comorbidomes changed. In the first instance, lowering the threshold resulted in the introduction of Hyperprolactinemia with an OR of approximately 15, Hepatitis C and Arthropathy. The entire list can be check in this table.

In the second instance, the new comorbidomes introduced were Hyperlipidemia and Arthropathy. We conclude that conditions like Arthropathy, Hyperprolactinemia and Hyperlipidemia seem to be quiet relevant in the mechanism of transmission. These same conditions will be present also in the following analyses of prevalences.

Matched Cohort Analysis: Age-Based Comorbidity Matching To improve the balance between groups and enhance the validity of our statistical comparisons, we implemented a **matching strategy** to pair **psoriatic patients** with and without infertility based on **age**. The goal was to create a more balanced dataset and refine the assessment of comorbidities associated with psoriasis and infertility.

Patient Matching Criteria The matching process was conducted as follows:

- **Cohort Selection:** All male patients diagnosed with psoriasis were considered.

- **Infertility Group:** Patients who developed infertility after psoriasis were identified.
- **Control Group:** Since more than 90% of psoriatic patients do not have an infertility diagnosis, a random sample of psoriatic patients without infertility was selected as controls.
- **Age Matching:** Each infertile patient was matched to a control patient of the same age at the time of infertility diagnosis.
- **Comorbidity Tracking:**
  - For infertility patients, only comorbidities developed between psoriasis and infertility diagnoses were considered.
  - For control patients, only comorbidities developed at or before the matched patient's infertility age were included.

This methodology ensures that thetime window for comorbidity accumulation is standardized, reducing potential bias due to differing disease durations.

3.7.4 Statistical Considerations

- The randomized matching ensures that the control group is age-balanced.
- The restriction on comorbidity timeframes provides a more accurate comparison between cases and controls.
- However, this method significantly reduces the sample size, potentially limiting the power to detect statistically significant associations.

3.7.5 Preliminary Results (INF $\rightarrow$ PSO)

The initial implementation of this method focused on the Infertility $\rightarrow$ Psoriasis (INF $\rightarrow$ PSO) analysis. The matched cohort size was 118 patients, with:

- 59 patients diagnosed with infertility before psoriasis.
- 59 matched controls (psoriatic patients without infertility).

Comorbidity Analysis Using logistic regression to compute odds ratios (ORs) for comorbidities in this matched dataset, no statistically significant associations were identified ($p < 0.05$ for all comorbidities). This lack of significance is likely due to:

- The **small sample size** (118 patients), which limits statistical power.
- The **stringent inclusion criteria for comorbidities**, as only conditions occurring within the defined time windows were considered.
- The **imbalance in comorbidity distributions**, where many patients lacked comorbidity diagnoses in the restricted timeframe.

3.7.6 Next Steps

Before extending this approach to the Psoriasis → Infertility (PSO → INF) analysis, the following refinements should be considered:

- Increasing the sample size by adjusting matching tolerances (e.g., allowing small age deviations).
- Exploring alternative statistical methods, such as Firth's logistic regression, which is better suited for small-sample binary outcomes.
- Reassessing the comorbidity inclusion criteria to prevent excessive data loss.

This matched cohort method provides a robust framework for balancing comparisons but requires further refinements to address the limitations imposed by sample size constraints.

3.7.7 Conclusion

- Excluding psoriatic patients without a diagnosis date revealed new significant comorbidities (Hyperprolactinemia, PVD) strongly associated with infertility.
- Control group selection remains a challenge, as non-infertile patients in the dataset may simply be too young to have developed or been diagnosed with infertility.

- Prostate cancer and BPH showed a bidirectional temporal relationship with psoriasis, indicating a potential interplay between inflammatory pathways and prostate disease.

Future research should:

- Investigate potential causal mechanisms linking psoriasis, infertility, and vascular disorders.
- Develop a better-matched control group that accounts for age and fertility status.
- Explore whether psoriasis treatment influences the onset of prostate disease.

4 Conclusions

The study investigated the complex relationship between psoriasis and male infertility, focusing on comorbidities, chronological disease progression, and predictive modeling. The results provide compelling evidence of a bidirectional link between the two conditions, with specific comorbid profiles emerging in affected patient groups.

4.1 Comorbidity Analysis in Psoriasis vs. Control Groups

Patients with psoriasis demonstrated significantly higher prevalence of several comorbidities compared to controls, reinforcing the systemic nature of the disease. Key findings include:

- **Metabolic and Cardiovascular:** Psoriasis patients exhibited higher rates of hyperlipidemia (46.1% vs. 42.4%, $p < 10^{-86}$), obesity (27.5% vs. 22.5%, $p < 10^{-208}$), diabetes (18.4% vs. 16.3%, $p < 10^{-22}$), and hypertension (29.3% vs. 26.5%, $p < 10^{-61}$).
- **Autoimmune and Rheumatic Diseases:** Increased prevalence of arthropathy (25.4% vs. 21.8%, $p < 10^{-109}$), Crohn's disease (0.72% vs. 0.40%, $p < 10^{-29}$), and ulcerative colitis (0.52% vs. 0.38%, $p < 10^{-9}$).

- **Neurological and Psychiatric Conditions:** Depression (8.5% vs. 7.5%, $p < 10^{-24}$) and anxiety (6.2% vs. 5.3%, $p < 10^{-13}$) were more common among psoriasis patients.
- **Pulmonary and Respiratory Disorders:** Higher rates of COPD (4.6% vs. 3.7%, $p < 10^{-34}$) and asthma (8.3% vs. 7.1%, $p < 10^{-29}$).
- **Oncological Risks:** Psoriasis patients had a higher prevalence of malignancies (11.2% vs. 10.2%, $p < 10^{-18}$), including non-Hodgkin's lymphoma (0.88% vs. 0.68%, $p < 10^{-7}$).

4.2 Psoriasis and Infertility: Clustering and Classification

Analysis of psoriasis and infertility patients revealed distinct comorbid profiles:

- Chi-squared tests identified significant conditions associated with infertility among psoriatic patients, including hyperthyroidism ($p = 1.83 \times 10^{-11}$), diabetes ($p = 4.07 \times 10^{-2}$), depression ($p = 1.29 \times 10^{-5}$), multiple sclerosis ($p = 2.35 \times 10^{-80}$), and joint replacement ($p = 6.42 \times 10^{-44}$).
- Clustering analysis using Gaussian Mixture Models (GMM) consistently identified two distinct patient subgroups, validated by silhouette scores around 0.8.
- Classification models, including Random Forest (AUC-ROC = 0.61) and StepMix (accuracy = 48.17%), provided moderate predictive accuracy, highlighting the complexity of the relationship.

4.3 Infertility and Psoriasis: Chronological Inversion

When examining patients who developed infertility before psoriasis, key findings included:

- Significant comorbidities associated with infertility included chronic renal failure ($p = 0.0193$), arthropathy ($p = 0.0072$), hyperthyroidism ($p = 0.0058$), diabetes ($p = 0.0246$), and COPD ($p = 0.0256$).

- Multiple Correspondence Analysis (MCA) confirmed two primary patient clusters, with optimal silhouette scores around 0.85 for 10 components.
- The StepMix model identified four latent classes (entropy = 0.51), with moderate classification accuracy (48.17%).

4.4 Comorbidome Analysis: Psoriasis and Male Infertility

The comorbidome analysis further clarified the relationship between psoriasis and infertility:

- Among psoriatic patients, infertile males exhibited lower prevalence of hyperlipidemia and hypertension (both $p < 0.001$).
- No significant comorbidities emerged when comparing infertile psoriatic patients to infertile controls, suggesting complex, multifactorial mechanisms underlying the association.

4.5 Refined Comorbidome Analysis: Excluding Patients Without Diagnosis Dates

Excluding psoriatic patients without a confirmed diagnosis date revealed additional significant findings:

- **Hyperprolactinemia (OR = 33.59)** and **Peripheral Vascular Disease (OR = 410.19)** were strongly associated with infertility among psoriatic patients, though wide confidence intervals highlighted the impact of rare conditions on statistical estimates.
- A bidirectional relationship emerged between psoriasis and prostate conditions, with prostate cancer and benign prostatic hyperplasia (BPH) showing temporal associations with psoriasis onset.

4.6 Matched Cohort Analysis: Age-Based Comorbidity Matching

To address potential confounding, age-based matching was applied:

- Among 118 matched patients (59 with infertility before psoriasis and 59 controls), no statistically significant associations were found, likely due to the small sample size and stringent inclusion criteria.

This comprehensive analysis confirms that psoriasis and male infertility are closely intertwined, with shared comorbidities, bidirectional risk factors, and complex disease pathways. The findings underscore the need for multidisciplinary care, further exploration of causal mechanisms, and improved cohort design for future research.

4.7 Final Summary and Next Steps

This study explored the complex relationship between psoriasis and male infertility, highlighting significant comorbidities, chronological patterns, and predictive factors. The findings underscore the multifaceted nature of these conditions and emphasize the importance of a holistic approach to patient care.
Key findings include:

- Psoriasis is strongly associated with metabolic, cardiovascular, autoimmune, neurological, and psychiatric comorbidities, reinforcing its systemic inflammatory nature.

- Infertility among psoriatic patients was linked to conditions such as hyperthyroidism, multiple sclerosis, and joint replacement, with clustering analysis identifying two distinct patient subgroups.

- The chronological inversion analysis revealed that patients who developed infertility prior to psoriasis exhibited higher rates of chronic renal failure, arthropathy, diabetes, and COPD.

- Comorbidome analysis identified hyperprolactinemia and peripheral vascular disease as particularly strong predictors of infertility among psoriatic patients, though the rarity of these conditions necessitates cautious interpretation.

- Age-based cohort matching, while useful for balancing comparisons, highlighted the limitations of small sample sizes and the need for more robust datasets.

The study's findings not only confirm a bidirectional link between psoriasis and male infertility but also highlight the importance of early detection and integrated care. Given the complex interplay of metabolic, autoimmune, and inflammatory pathways, future research should focus on:

- Expanding datasets to include larger cohorts and more documented infertility cases for improved statistical power.
- Investigating causal mechanisms through longitudinal studies, including genetic and inflammatory biomarkers.
- Refining predictive models by incorporating age, BMI, medication use, and lifestyle factors.
- Exploring the impact of psoriasis treatments, particularly IL-23 inhibitors, on fertility outcomes.

Ultimately, this study reinforces the need for multidisciplinary collaboration between dermatologists, urologists, endocrinologists, and data scientists to develop personalized treatment strategies and improve patient outcomes. This will be further explored by means of clustering of the 122 comorbidities in a more compact number of larger categories to further explore ORs and prevalences and convalidate these results with a larger and more diversified dataset from TriNexT. We believe that this new dataset with more than 1000 male infertile patients will give us more inference power and the diverse backgrounds of the sample points will enhance the results, which, up until now, were skewed towards a precise ethnicity and medical practices pertaining the State of Israel.

5 Tables and Images

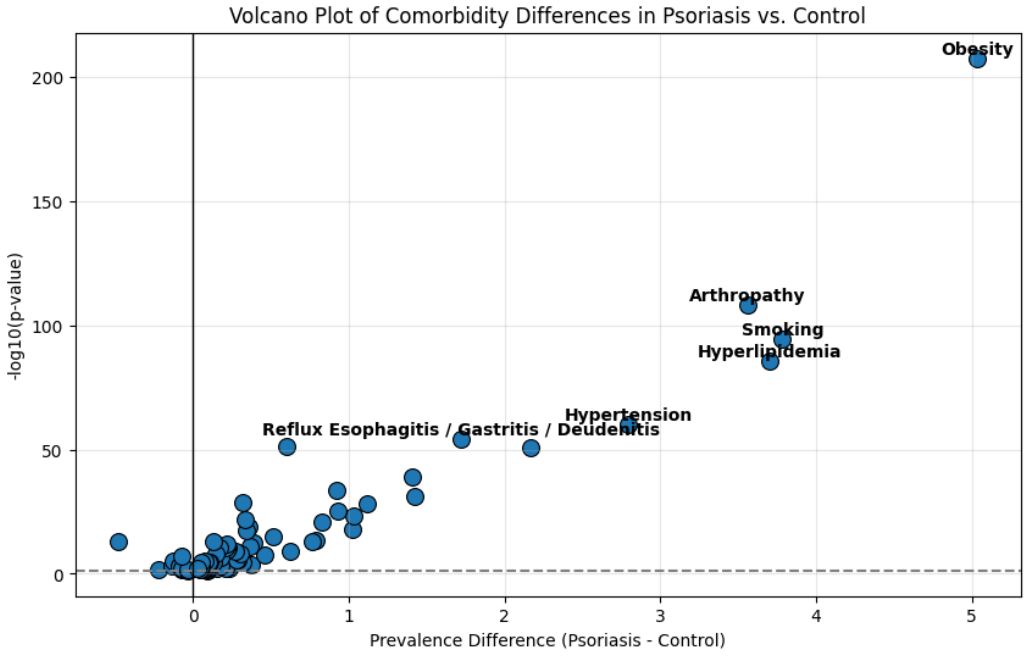

Figure 9: Prevalence Differences between Psoriasis versus Control for each Comorbidity.

	Comorbidity	Psoriasis Prevalence (%)	Psoriasis CI (%)	Control Prevalence (%)	Control CI (%)	Chi2	p-value	dof
13	Smoking	30.632357	[30.29, 30.98]	0.115603	[0.08, 0.16]	11683.956862	0.000000e+00	1
8	Hyperlipidemia	24.703560	[24.38, 25.03]	0.102391	[0.07, 0.15]	8943.710998	0.000000e+00	1
7	Obesity	15.469621	[15.20, 15.74]	0.052847	[0.03, 0.09]	5204.731206	0.000000e+00	1
33	Arthropathy	14.553287	[14.29, 14.82]	0.062756	[0.04, 0.10]	4851.099129	0.000000e+00	1
22	Hypertension	12.534759	[12.29, 12.78]	0.066059	[0.04, 0.10]	4104.107285	0.000000e+00	1
25	Asthma	7.068451	[6.88, 7.26]	0.006606	[0.00, 0.02]	2244.239251	0.000000e+00	1
5	Diabetes	6.822078	[6.64, 7.01]	0.049544	[0.03, 0.08]	2129.169463	0.000000e+00	1
27	Reflux Esophagitis / Gastritis / Deudenitis	4.888556	[4.73, 5.05]	0.023121	[0.01, 0.05]	1513.262278	0.000000e+00	1
4	Hypothyroidism	4.699814	[4.54, 4.86]	0.026424	[0.01, 0.05]	1449.648577	0.000000e+00	1
18	IHD	3.976544	[3.83, 4.12]	0.026424	[0.01, 0.05]	1216.927728	1.277771e-266	1
2	Malignancy	3.694152	[3.56, 3.84]	0.026424	[0.01, 0.05]	1126.719761	5.147071e-247	1
31	other kidney Disease	3.656692	[3.52, 3.80]	0.016515	[0.01, 0.04]	1122.166408	5.025576e-246	1
11	Depression	3.393030	[3.26, 3.53]	0.019818	[0.01, 0.04]	1035.838694	2.913139e-227	1
35	Osteoporosis	2.950711	[2.83, 3.08]	0.016515	[0.01, 0.04]	898.310653	2.285941e-197	1
15	Other Neurological Disease	2.723068	[2.60, 2.85]	0.013212	[0.01, 0.03]	829.057954	2.597037e-182	1
12	Anxiety	2.715864	[2.60, 2.84]	0.013212	[0.01, 0.03]	826.792525	8.072396e-182	1
26	Peptic Ulcer	2.257697	[2.15, 2.37]	0.009909	[0.00, 0.03]	685.631148	3.984296e-151	1
29	Prostatic Hypertrophy	2.166928	[2.06, 2.28]	0.009909	[0.00, 0.03]	657.289700	5.804416e-145	1
21	Arrhythmia	2.030055	[1.93, 2.14]	0.016515	[0.01, 0.04]	609.752556	1.266394e-134	1
17	Glaucoma	1.482559	[1.40, 1.58]	0.013212	[0.01, 0.03]	442.355394	3.325253e-98	1
24	COPD	1.321192	[1.24, 1.41]	0.003303	[0.00, 0.02]	399.825448	6.010804e-89	1
23	s/p CVA	1.260680	[1.18, 1.35]	0.009909	[0.00, 0.03]	376.319130	7.876888e-84	1
16	Retinopathy	1.257798	[1.18, 1.34]	0.009909	[0.00, 0.03]	375.431511	1.229161e-83	1
10	Neuroses	1.234746	[1.16, 1.32]	0.013212	[0.01, 0.03]	365.922409	1.445524e-81	1
30	Chronic Renal Failure	1.162707	[1.09, 1.25]	0.006606	[0.00, 0.02]	348.578018	8.645661e-78	1
39	OncBreat	0.847177	[0.78, 0.92]	0.006606	[0.00, 0.02]	251.722597	1.093712e-56	1
9	Psychoses	0.717507	[0.66, 0.78]	0.006606	[0.00, 0.02]	212.046467	4.913544e-48	1
38	Congenital Anomalies	0.688691	[0.63, 0.75]	0.003303	[0.00, 0.02]	205.646161	1.224042e-46	1
19	CHF	0.668520	[0.61, 0.73]	0.003303	[0.00, 0.02]	199.482260	2.709038e-45	1
3	Hyperthyroidism	0.654113	[0.60, 0.72]	0.003303	[0.00, 0.02]	195.080564	2.474164e-44	1
36	Joint Replacement	0.554699	[0.50, 0.61]	0.003303	[0.00, 0.02]	164.733736	1.045863e-37	1
37	Gout	0.482660	[0.43, 0.54]	0.003303	[0.00, 0.02]	142.770699	6.597309e-33	1
34	Rheumatoid Arthritis	0.476897	[0.43, 0.53]	0.003303	[0.00, 0.02]	141.014664	1.597113e-32	1
0	Hepatitis B Carrier	0.446641	[0.40, 0.50]	0.003303	[0.00, 0.02]	131.797964	1.656482e-30	1
40	OncColon	0.344346	[0.30, 0.39]	0.003303	[0.00, 0.02]	100.668380	1.087478e-23	1
41	OncProatate	0.289596	[0.25, 0.33]	0.003303	[0.00, 0.02]	84.028660	4.876539e-20	1
1	Familial Mediteranean Fever	0.286715	[0.25, 0.33]	0.003303	[0.00, 0.02]	83.153325	7.592953e-20	1
28	Chronic Act/Per Hepatitis	0.252136	[0.22, 0.29]	0.003303	[0.00, 0.02]	72.652989	1.545726e-17	1
20	Cardiomyopathy	0.230524	[0.20, 0.27]	0.003303	[0.00, 0.02]	66.093951	4.299332e-16	1
43	OncUrineBlader	0.211794	[0.18, 0.25]	0.006606	[0.00, 0.02]	58.099377	2.492052e-14	1
6	Hypo/Hyperparathyroidism	0.185860	[0.16, 0.22]	0.003303	[0.00, 0.02]	52.548993	4.196478e-13	1
45	OncOther	0.135433	[0.11, 0.17]	0.003303	[0.00, 0.02]	37.278806	1.023913e-09	1
42	OncLung	0.092210	[0.07, 0.12]	0.003303	[0.00, 0.02]	24.224477	8.573573e-07	1
44	OncOral	0.057631	[0.04, 0.08]	0.003303	[0.00, 0.02]	13.841159	1.989305e-04	1
14	Motor Neuron Disease	0.050427	[0.04, 0.07]	0.003303	[0.00, 0.02]	11.694683	6.267894e-04	1

Figure 10: Prevalence Differences between Psoriasis versus Control for each Comorbidity. Psoriasis Diagnosis After Comorbidity Diagnosis.

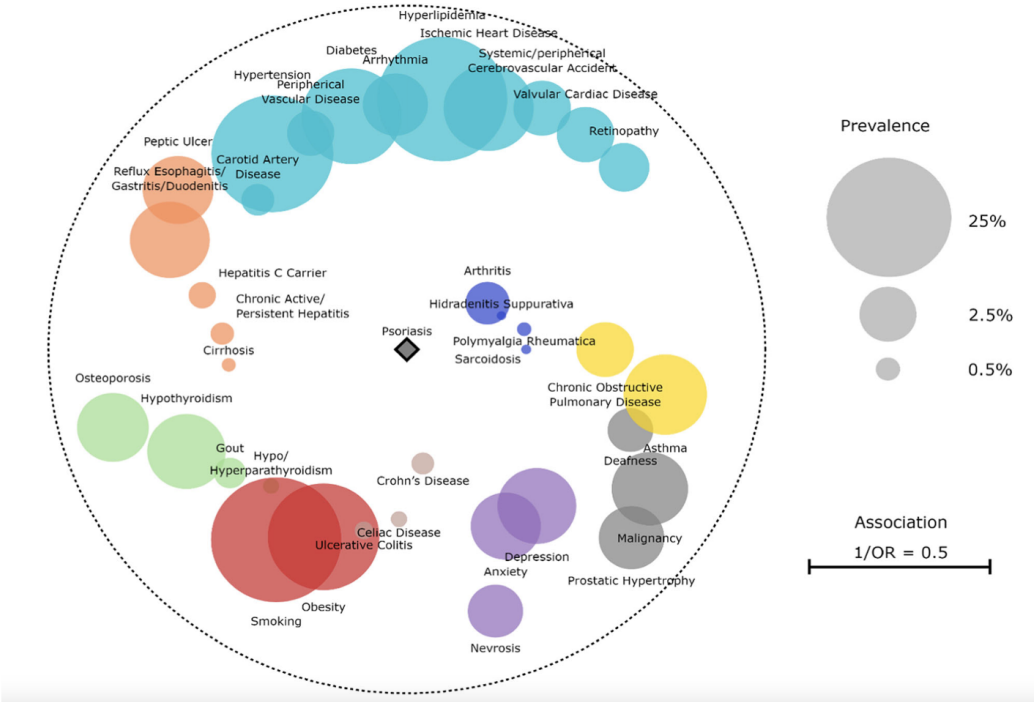

Figure 11: Comorbidome Psoriasis versus Control Group. Courtesy of Giovanni Damiani (2023).

References

[1] Damiani, G. (2023). Fertilità maschile e IL-23 inibitori nei pazienti con psoriasi moderata-severa: dalla valutazione dei parametri spermatici alla costruzione di un piano diagnostico terapeutico ed assistenziale (PDTA).

[2] Damiani, G., Bragazzi, N. L., Karimkhani Aksut, C., Wu, D., Alicandro, G., McGonagle, D., Guo, C., Dellavalle, R., Grada, A., Wong, P., La Vecchia, C., Tam, L. S., Cooper, K. D., Naghavi, M. (2021). The Global, Regional, and National Burden of Psoriasis: Results and Insights From the Global Burden of Disease 2019 Study. *Frontiers in Medicine (Lausanne)*, 8, 743180.

[3] Danielsen, K., Olsen, A. O., Wilsgaard, T., Furberg, A. S. (2013). Is the prevalence of psoriasis increasing? A 30-year follow-up of a population-based cohort. *British Journal of Dermatology*, 168, 1303–1310.

[4] Kimball, A. B., Guerin, A., Latremouille-Viau, D., Yu, A. P., Gupta, S., Bao, Y., Mulani, P. (2010). Coronary heart disease and stroke risk in patients with psoriasis: Retrospective analysis. *American Journal of Medicine*, 123, 350–357.

[5] Johansen, C., Usher, P. A., Kjellerup, R. B., Lundsgaard, D., Iversen, L., Kragballe, K. (2009). Characterization of the interleukin-17 isoforms and receptors in lesional psoriatic skin. *British Journal of Dermatology*, 160(2), 319–324.

[6] Nograles, K. E., Zaba, L. C., Guttman-Yassky, E., Fuentes-Duculan, J., Suárez-Fariñas, M., Cardinale, I., Khatcherian, A., Gonzalez, J., Pierson, K. C., White, T. R., Pensabene, C., Coats, I., Novitskaya, I., Lowes, M. A., Krueger, J. G. (2008). Th17 cytokines interleukin (IL)-17 and IL-22 modulate distinct inflammatory and keratinocyte-response pathways. *British Journal of Dermatology*, 159(5), 1092–1102.

[7] Hartupee, J., Liu, C., Novotny, M., Li, X., Hamilton, T. (2007). IL-17 enhances chemokine gene expression through mRNA stabilization. *Journal of Immunology*, 179(6), 4135–4141.

[8] Quaderni del Ministero della Salute. (2012). Criteri di appropriatezza strutturale, tecnologica e clinica nella prevenzione, diagnosi e cura delle patologie andrologiche.

[9] Naz, M., Kamal, M. (2017). Classification, causes, diagnosis, and treatment of male infertility: A review. *Oriental Pharmacy and Experimental Medicine*, 17, 89–109.

[10] Griswold, M. D. (1998). The central role of Sertoli cells in spermatogenesis. *Seminars in Cell and Developmental Biology*, 9, 411–416.

[11] Tsigos, C., Papanicolaou, D. A., Kyrou, I., Raptis, S. A., Chrousos, G. P. (1999). Dose-dependent effects of recombinant human interleukin-6 on the pituitary-testicular axis. *Journal of Interferon and Cytokine Research*, 19, 1271–1276.

[12] Ding, L., Linsley, P. S., Huang, L. Y., Germain, R. N., Shevach, E. M. (1993). IL-10 inhibits macrophage costimulatory activity by selectively inhibiting the up-regulation of B7 expression. *Journal of Immunology*, 151, 1224–1234.

[13] Pentikäinen, V., Erkkilä, K., Suomalainen, L., Otala, M., Pentikäinen, M. O., Parvinen, M., Dunkel, L. (2001). TNF-alpha down-regulates the Fas ligand and inhibits germ cell apoptosis in the human testis. *Journal of Clinical Endocrinology and Metabolism*, 86, 4480–4488.

[14] Janitz, M., Fiszer, D., Michalczak-Janitz, K., Lukaszyk, A., Fernandez, N., Skorupski, W., Kurpisz, M. (1994). Analysis of mRNA for class I HLA on human gametogenic cells. *Molecular Reproduction and Development*, 38, 231–237.

[15] Singh, S., Kroe-Barrett, R. R., Canada, K. A., Zhu, X., Sepulveda, E., Wu, H., et al. (2015). Selective targeting of the IL-23 pathway: Generation and characterization of a novel high-affinity humanized anti-IL-23A antibody. *MAbs*, 7(4), 778–791.

[16] Presta, L. G. (2012). Engineered anti-IL-23P19 antibodies. U.S. Patent No. 8,293,883 B2.

[17] Zhou, L., Wang, Y., Wan, Q., Wu, F., Barbon, J., Dunstan, R., et al. (2021). A non-clinical comparative study of IL-23 antibodies in psoriasis. *MAbs*, 13(1), 1964420.

[18] Luo, J., Wu, S. J., Lacy, E. R., Orlovsky, Y., Baker, A., Teplyakov, A., et al. (2010). Structural basis for the dual recognition of IL-12 and IL-23 by ustekinumab. *Journal of Molecular Biology*, 402(5), 797–812.

[19] Beyer, B. M., Ingram, R., Ramanathan, L., Reichert, P., Le, H. V., Madison, V., et al. (2008). Crystal structures of the pro-inflammatory cytokine interleukin-23 and its complex with a high-affinity neutralizing antibody. *Journal of Molecular Biology*, 382(4), 942–955.

[20] Pettersen, E. F., Goddard, T. D., Huang, C. C., Couch, G. S., Greenblatt, D. M., Meng, E. C., et al. (2004). UCSF Chimera–a visualization system for exploratory research and analysis. *Journal of Computational Chemistry*, 25(13), 1605–1612.

[21] Armstrong, A. W., Puig, L., Joshi, A., Skup, M., Williams, D., Li, J., et al. (2020). Comparison of Biologics and Oral Treatments for Plaque Psoriasis: A Meta-analysis. *JAMA Dermatology*, 156(3), 258–269.