Published in final edited form as:

Clin Exp Dermatol. 2019 June ; 44(4): e151–e152. doi:10.1111/ced.13954.

Hidradenitis Suppurativa and Diabetes: Big Data Bias Masks a True Association

DR J.W. Frew

Laboratory of Investigative Dermatology, The Rockefeller University, NY, USA,
jfrew@rockefeller.edu

Keywords

Hidradenitis Suppurativa; Diabetes Mellitus; Meta-Analysis; Bias

Phan et al¹ present an updated meta-analysis of the association between Hidradenitis Suppurativa (HS) and Type 2 Diabetes Mellitus (DM) which demonstrates a statistically significant elevation in the odds ratio (OR) for DM in HS Patients compared with healthy controls (OR=2.17 95% CI 1.85–2.55 $p<0.001$)¹. They astutely highlight that the risk difference of 0.4% between HS patients and controls (16.1% vs 15.7%) implies that this result is clinically non-significant, quite at odds with clinical experience, data from existing studies and the remainder of the article discussing the importance of DM awareness in HS patients and possible mechanistic links between DM and HS.

The reason for this statistical-common sense disconnect is the lack of acknowledgement of external bias in Phan's analysis. Whilst the methodology is sound, it does not eliminate all bias and residual confounding despite the high reported I^2 (which as a point estimate is misleading in the setting of large heterogeneity)². Although weighting of studies addresses 'within study variance' and heterogeneity (I^2) assesses methodological (internal) heterogeneity^{2,3}, no consideration of external biases (including population bias and outcome bias) is provided by a statistical program, it requires the use of common sense on behalf of the statistician and authors^{2,3} often by plotting and visualising the data.

We replicated Phan et al's analysis¹ using the same methodology, with the additional step of assessing studies for external bias and threats to generalisability^{2,4}. Visualisation of data in a L'Abbe plot (Figure 1) demonstrates lack of clustering on the x axis- indicating violation of the assumption of one underlying baseline risk^{2,4}. Examining risks to external bias in the original analysis¹, four population-based studies (Garg, Kirby, Gold and Shalom)¹ utilise claims-based insurance data. Concerns have been raised (particularly in USA claims-based data analyses) that such data is not indicative of the general population⁵. Such data may suffer from selection bias due to the general population's variable access to healthcare, insurance, as well as issues regarding 'over diagnosis, over-coding' (outcomes bias) in these datasets⁵. Two of these identified studies are outliers in the L'Abbe plot.

Conflicts of Interest: The author has no conflicts of interest to declare

Removing these points and reconducting the analysis increases the I^2 to 96% (from 86%) and revises the OR to 2.95 (1.76–4.94) (Figure 2) and the risk difference of 3.85% (HS= 8.55% to Controls= 4.69%). This replication addresses the issues of heterogeneity in Phan's initial analysis. It also raises the vital issue that meta-analysis of observational studies is at high risk for external bias. It also highlights the importance of intelligent consideration of the external validity of results and not merely relying upon the results of statistical program outputs.

Our replication analysis concludes that, for the populations included in this analysis, there is a clinically and statistically significant elevation in the risk of Diabetes Mellitus in patients with Hidradenitis Suppurativa. The use of population scale data from claims-based data has the suggestion of a different risk population compared to those recruited from prospective and retrospective case-control trials. Such differences need to be considered in the generalisability of observational studies to clinical situations.

Acknowledgments

Funding Sources: Supported in part by grant # UL1 TR001866 from the National Center for Advancing Translational Sciences (NCATS), National Institutes of Health (NIH) Clinical and Translational Science Award (CTSA) program.

References:

- 1). Phan K, Charlton O, Smith SD Hidradenitis Suppurativa and Diabetes Mellitus: updated systematic review and adjusted meta-analysis. *Clin Exp Dermatol* 2019 Doi:10.1111/ced.13922
- 2). Kaplan RM, Chambers DA, Glasgow RE Big Data and Large Sample Size: A Cautionary Note on the Potential for Bias. *Clin Transl Sci*, 7: 342–346 [PubMed: 25043853]
- 3). Von Hippel PT The heterogeneity statistic I^2 can be biased in small meta-analyses. *BMC Med Res Methodol* 2015;15:35 [PubMed: 25880989]
- 4). Bax L, Kleda N, Fukui N, Yaju Y Tsuruta H, Moons KGM More than Numbers: The Power of Graphs in Meta Analysis. *Am J Epidemiol* 2009;169(2):249–255 [PubMed: 19064649]
- 5). Jensen ET, Cook SF, Allen JK, Logie J, Brookhart MA, Kappelman MD, Dellon ES Enrollment factors and bias of disease prevalence estimates in administrative claims data. *Ann Epidemiol* 2015;25(7):519–525 [PubMed: 25890796]

Figure 1:

L'Abbe plot of studies included in Phan et al's meta-analysis of observational studies evaluating the association between Hidradenitis Suppurativa (HS) and Type 2 Diabetes Mellitus (T2DM). Red symbols indicate studies with identified external bias which were removed; and blue symbols indicate studies which were retained in the replicated meta-analysis. The central black line indicates equivalency (the same proportion of patients with T2DM in the HS and healthy control groups). All studies demonstrated a positive association of T2DM with HS, with spread along the vertical axis consistent with possible variation in T2DM prevalence with HS disease severity. Spread along the horizontal axis indicates variable baseline risk of T2DM in the healthy control population suggestive of external bias and other threats to generalisability.

Figure 2:
 Forrest Plot of Replicated Meta-Analysis after removal of the four studies identified to be at high risk of external bias.