

The following papers have looked directly or indirectly at data quality of THIN, validity of diagnoses and reproducibility:

Lewis JD, Schinnar R, Bilker WB, Wang X, Strom BL. Validation studies of the health improvement network (THIN) database for pharmacoepidemiology research. *Pharmacoepidemiol Drug Saf* 2007; 16: 393–401.

- Comment: They state “In conclusion, we have successfully reproduced a large number of well-established epidemiological associations in the new THIN database. In addition, we have documented similar results using data from the THIN practices that have or continue to participate in GPRD* and those that never participated in GPRD. These data support the validity of the data within THIN for use in pharmacoepidemiology research.”

Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Brindle P. Performance of the QRISK cardiovascular risk prediction algorithm in an independent UK sample of patients from general practice: a validation study. *Heart* 2008;94;34-39

- The authors used a THIN cohort to validate the QRisk score developed in QResearch.
- They state: “Table 2 shows the completeness of the recording for each risk factor for men and women in both cohorts [from THIN and QResearch]... Overall, levels of recording were very similar between the two cohorts”

Garcia Rodriguez LA, Barreales Tolosa, L. Risk of upper gastrointestinal complication among users of traditional NSAIDs and COXIBs in the General population. *Gastroenterology* 2007;132:498–506

- The authors showed high diagnostic validity of Upper Gastrointestinal complications (UGIC). A total of 96% were confirmed by the GP and the site of bleeding/perforation was 100%.
- They observed similarity in terms of detection of outcomes between THIN only practices and those that had at sometime contributed to GPRD, “we assessed the risk of UGIC among current users of NA-NSAIDs according to whether the practice was also a contributor to the GPRD or not and found no major difference in estimates of risk between the 2 subsets” [see page 501]

Arellano FM, Conde, E, Wenworth CE, Schlienger R, Fernandez-Vidaurre C, Arana A. Validation of cases of lymphoma in THIN. *Pharmacoepidemiol Drug Saf*, 2008; 17: S87–88 [ICPE abstract]

- A sample of 742 cases of lymphoma via Read codes were checked for validity by THIN practices; the positive predictive value was 92%. The authors state that “Validation of cases lymphoma in THIN yielded a high positive predictive

value and allowed classification into NHL [non-Hodgkin lymphoma] or HD [Hodgkin's Disease] for most cases”

Van Staa TP, Geusens P, Pols HAP, de Laet C, Leufkens HGM, Cooper C. A simple score for estimating the long-term risk of fracture in patients using oral glucocorticoids. *Q J Med* 2005; 98:191–198

- THIN was used to validate a predictive model created in the GPRD.
- They state “When applying the risk score as developed in GPRD to another dataset (THIN), the risk score equally differentiated between high- and low-risk patients” Therefore, THIN reproduced results found in GPRD; this does not guarantee completeness but it means that both sources had similar ability to detect cases (fractures).