Predicting claims in life and especially disability insurance has, until now, been a very imprecise business. While the average life expectancy of a person based on gender, age and occupation are well documented in many countries, dealing with the current and previous medical conditions of applicants throws up enormous problems. The core problem lies in the relatively low number of claims in comparison to the overwhelming number of relevant medical conditions and their combinations. This is particularly true for disability insurance.

For example, the leading life insurers in Germany have around 1,000 (one thousand) disability claims per year. After segmenting the applicant base into two genders, around five age groups and around ten occupational categories, they are left with on average around ten claims per year per segment.

In the next step, each applicant may name any of around 16 thousand current or previous medical conditions (as represented by the ICD-10 coding system), one of 256 million imaginable combinations of two conditions, or maybe one of the 4 trillion \((4\times10^{12})\) combinations of three conditions. Clearly the actuary with ten claims per year for each segment has no means of modelling available which can lead to significant results in this situation.

The solution to this problem lies in the introduction of an intermediate step which we call "serious illness". Most death and disability claimants were relatively healthy at application time, but have transitioned through serious illness on the way to becoming disabled or dying.

The progression from known "serious illness" to death and disability is well documented in many countries through insurers, pension authorities and other government institutions. The step from unthreatening medical conditions to serious illness is thus the missing piece. This information, however, is available in the databases of health insurers, and is available in sufficient quantities to allow accurate statistical modelling.

Thus the complete chain from typical applicants with relatively unthreatening medical histories, to disability and death can be modelled with statistically significant results, allowing life insurers to differentiate their offers with a previously unknown degree of accuracy.

Keywords: Multivariate modelling, disability risk, mortality risk, morbidity risk.