Outlier identification: Why and how are funnel plots derived?

As part of the National CKD Audit in Primary Care (NCKDA) we need to identify and contact practices where rates CKD coding are observed to be significantly lower than average. The audit aims to identify outlying practices with regards to two key outcomes of interest: (i) practice prevalence of coded CKD 3-5 (patients on the CKD register) and (ii) prevalence of uncoded CKD amongst those with eGFR evidence of CKD.

What are Funnel plots used for?

Funnel plots are a statistical method to identify outlying practices. Funnel plots allow viewers to visualise variation according to practice size and other features, and control lines are used to identify outliers who are 2 standard deviations (SDs) outside of the practice wide variation, or 3 SDs outside the practice wide variation. Wide variation between practices in terms of both testing for and coding of CKD results in considerable statistical 'overdispersion'. This phenomenon has been described previously in between-centre comparisons of similar types of healthcare data\(^1\). Some of this overdispersion can be accounted for by adjusting the control lines for between-practice factors. We have opted to apply multiplicative random-effects methods in the NCKDA as this approach is more conservative with respect to practices with very small numbers of expected cases. More details on how we manage overdispersion is presented in the appendix.

Data

Detailed information on the which, and how, data are collected are available on the NCKDA website (http://www.ckdaudit.org.uk/audit/what-information-collected/#.VmrwJRHTa9d). Briefly data are collected using BMJ-Informatica software from GP practices who participate in the National Chronic Kidney Disease Audit (NCKDA) in England and Wales. The Informatica software extracts information on all individuals within a practice aged 18+ from 2008 onwards who have either a QOF code for CKD (any stage; see website for list of codes), a creatinine measurement, or with a risk factor / renal disease diagnosis at least one year prior to data extract. Retrospective creatinine and estimated glomerular filtration rate (eGFR) laboratory reports were also obtained for all these individuals, together with baseline characteristics including age, sex and postcode (to derive an index of multiple deprivation).

If approved by the practice, complete practice age-sex distribution data (in 5-year age bands) were also obtained or otherwise obtained from publicly available practice demographics.\(^2\)

Methodology

The two key outcome measures for outlier identification were determined as follows:

(i) Practice coded CKD prevalence or ‘observed CKD prevalence’

Practice coded CKD prevalence is calculated by identifying total patient numbers in a practice with codes for CKD stages 3-5. Direct age/sex standardisation is carried out using list-size data of the GP practice for all ages (using 5-year sex-specific age-bands). Note this calculation assumes no CKD in under 18s (as these are not observed here), although this is not unreasonable as such these are very small numbers or even unobserved in some practices.

(ii) Uncoded CKD amongst those with eGFR evidence of CKD

Patients with eGFR evidence of CKD are defined as those with either: (a) two most recent eGFR measurements are both <60ml/min and where at least 3-months have elapsed between measurements (the most recent measurement must be in the last 2 years and both measurements since 1/1/2008), or (b) the most recent eGFR measurement (since 1/1/2008, in last 2 years) is <60ml/min, and this is the only eGFR measurement ever taken. Individuals meeting one of these criteria are then defined as having uncoded CKD if they do not have a code for stage 3-5 CKD.
(i) Identifying practices with an outlying low coded CKD 3-5 prevalence

Target CKD prevalences for each practice should ideally reflect underlying differences in the practice populations, and the acceptable range of values should also reflect the population size from which the observed value is derived. This can be achieved using a funnel plot based on the number of observed CKD cases divided by the number of expected cases plotted against the number of expected cases.

Calculating expected number of coded CKD cases

The number of expected cases can be estimated from an adjusted model:

\[
\text{logit}\left(\pi_j\right) = \beta_0 + \beta_1\text{diabetes}_j + \beta_2\text{ht}_j + \beta_3\text{cvd}_j + \beta_4\text{black}_j + \beta_5\text{imd}_j
\]

where diabetes, ht, cvd, and black give the practice-level proportions of individuals with diabetes, hypertension, CVD and black ethnicity respectively, and imd gives with practice median IMD rank (calculated from individual IMD ranks for those with CKD or at increased risk of CKD only). \(j\) gives the proportion of coded CKD cases out of the practice list-size, standardised for age and sex (numbers of cases are back-calculated after standardisation in order to do this).

Figure 1. Funnel plot of age-standardised CKD prevalence by practice

The funnel plot displays the number of observed CKD cases divided by the expected cases on the Y-axis and the number of expected cases on the X-axis. The expected number of cases is calculated from the practice size and demographics (age, sex, diabetes, hypertension, CVD, black ethnicity and deprivation). A practice lying on the line \(Y=1\) has a coded CKD prevalence performing close to average, whereas a practice lying above \(Y=1\) has a coded CKD prevalence that is higher than the average. For CKD prevalence, outliers are those practices who fall below the lower 2 and 3 SD control lines i.e. these are practices coding far fewer CKD cases than would be expected given their practice size and demographic profile.
(iii) Identifying practices with an outlying high proportion of uncoded CKD patients amongst those with biochemical evidence for CKD

For this measure the proportion of uncoded CKD patients out of the number of individuals with GFR evidence of CKD for outcome is used. This is then adjusted for the proportion of the practice population who are known to have black ethnicity.

Figure 2. Funnel plot of proportion of uncoded CKD in those with GFR evidence of CKD by practice. Expected cases adjusted for proportion of practice with black ethnicity.

The funnel plot shows a ratio of observed cases of uncoded CKD divided by the number of expected cases by practice on the Y-axis and the expected number of uncoded CKD cases (adjusted the proportion of people with black ethnicity) on the X axis. A practice lying at Y=1 has an average proportion patients who have eGFR evidence of CKD but are uncoded for CKD. A practice lying above Y=1 has more than an average number of patients who are uncoded for CKD where there is biochemical evidence for CKD. For uncoded CKD, outlying practices are those that fall above the upper 2 and 3 SD control lines. i.e. practices with a large proportion of patients with biochemical CKD who remain uncoded for CKD.
Appendix:

Adjustment for overdispersion

Unadjusted (unadjusted for overdispersion, but adjusted for baseline characteristics) z-scores are first calculated for each practice, using the “observed” number of cases back-calculated from the age-sex standardised prevalence, and the expected number of cases using the model described in [1]. Since we are using an O/E outcome, a square-root transformation is applied (see reference 1) such that \( y = \sqrt{O/E} \), the standard \( t = 1 \) and \( S_0 = 1/2 \sqrt{E} \), and then:

\[
z = 2\sqrt{E}(y - 1) = 2(\sqrt{O} - \sqrt{E})
\]

In order to formally account for overdispersion using the multiplicative random-effects modelling, the control limits are adjusted to be \( t \pm ks_0\sqrt{\hat{\phi}} \), where \( \hat{\phi} = \sum_i z_i^2 / l \) is the number of practices, and \( k \) is the multiplier for \( s_0 \) for each contour (here, taking values 2 and 3). These values are then back-transformed (squared) to obtain values on the original O/E scale. Where the estimate of \( \hat{\phi} \) is high (i.e. there is large between-practice heterogeneity), some pre-back-transformation values may be negative, resulting in non-monotonically changing contours at low values on the x-axis. This has been addressed here by setting such post-transformation values to zero.

Winsorisation

This process further reduces the potential influence of extreme values on the contours whilst still retaining the same number of z-scores. This is achieved by ranking the z-values calculated above, and replacing the bottom \( w\% \) with the value of the \( w\th \) centile and top \( w\% \) with the value of the \( (100-w)\th \) centile, for \( w\% \) winsorisation. In the reporting for the NCKDA a 10% winsorisation is used throughout.

References

2 (HSCIC) HSCIC. Numbers of Patients Registered at a GP Practice - January 2015. 2015