Introduction:
A clinical database should be representative of the labelled population and guarantee completeness and accuracy of collected data. Without explicit permission of the patients, Swiss laws regarding data protection do not allow external audits based on periodic checks of random samples, supposed to give a general pattern of accuracy. To test alternative methods for quality control we introduced the principles of statistical process control to derive funnel plots from the Swiss ICU – Minimal Data Set (MDSi).

Methods:
The MDSi from all certified adult Swiss ICUs (2014 and 2015) was subjected to quality assessment (completeness and accuracy). For the analysis of accuracy, a list of logical rules and cross-checks was developed as e.g. range of SAPS II according to age. Errors were classified in coding errors (e.g. NEMS score > 56 points) or implausible data (NEMS without basic monitoring). We also checked for ICUs producing significantly more errors - outliers (> mean ± 3 standard deviations [SD] or > 99.8% confidence interval [CI] of an adapted version of the funnel plots, which allows the presence of trends depending of the ICU’s size.

Results:
A total of 164’415 patient MDSi (31 items/patient; 32 items for trauma patients) from the 77 certified ICUs were investigated. We detected 15’572 patients (9.5%) with an overall sum of 3121 coding errors and 31’265 implausible situations. Implausible situations related to supposedly inaccurate definitions (diagnostic and patient’s provenance prior to ICU admission) and discrepancies in the logical rules between diagnostics and treatments. Figure 1 is an example for imprecise coding of the diagnostic: 11 ICUs declared having treated 14-61% of their patients without a defined diagnosis.

Conclusion:
Accuracy of data in MDSi needs further improvement. Funnel plots may be useful for meaningful interpretation of data quality and permit to identify ICUs disproportionately generating inaccurate and/or implausible data.
Funnel plot depicting for each ICU its relation between the number of treated patients (x-axis) and the percentage of these patients with imprecise coding of the diagnosis (y-axis). The red (dotted) lines denote the 99.8% (95%) Confidence Intervall. The central line represents "local medians" estimated by the Locally Weighted Scatterplot Smoother – method.