In the present work we study behaviour of tools intended for code review and how they aim at eliminating security vulnerabilities. There is a lot of such tools, but a smaller set of them suffice to effectively improve the security of software. We provide results of empirical testing of these tools on artificial data in order to map vulnerability classes they are able to identify and also on real data in order to test their scalability.