Классификация биомаркеров: маркеры-диагносты и маркеры-классификаторы

Рубанович А.В., Сальникова Л.Е.

ФГБУН «Институт общей генетики им. Н.И. Вавилова РАН», Москва

Степень сопряженности биомаркера с бинарным фенотипическим признаком принято характеризовать отношением шансов (OR). Известно, что высокие значения этого показателя далеко не всегда гарантируют классификационную и/или диагностическую эффективность теста (Poste, 2011; Kraemer et al., 2011; Pepe et al., 2010; Kraft et al., 2009; Jakobsdottir et al, 2009), которые характеризуются разностью абсолютных рисков: SE+SP-1 (в исследованиях «случаи - контроли») и PPV+NPV-1 (в когортных исследованиях). Здесь использованы стандартные обозначения: SE=P(M|D) и SP = P(не M|не D) - чувствительность и специфичность теста; PPV=P(D|M) и NPV=P(не D|не M) - диагностическая эффективность положительных и отрицательных результатов тестирования, соответственно; M означает носительство маркера; D - наличие заболевания.

В настоящей работе показано, что предельное поведение показателей эффективности маркера при OR >> 1 существенно зависит от соотношения между частотами встречаемости маркера (pM) и распространенности заболевания (pD). Возможны две альтернативные ситуации: 1) при OR >>1 и pM >> pD показатель PPV+NPV-1 заведомо мал; 2) при OR >>1 и pM << pD заведомо малы чувствительность теста и соответствующий показатель эффективности SE+SP-1. Для этих двух ситуаций имеет смысл ввести специальные термины, отражающие специфику маркера.

При высокой частоте встречаемости маркера (pM >> pD) и OR >> 1 имеем высокую чувствительность, но низкую диагностическую значимость теста (PPV = pD/pM <<1). Маркер рационально использовать для массового скрининга и профессионального отбора. По результатам тестирования можно отобрать группу заведомо здоровых людей (свободных от маркера). При этом для носителей маркера вероятность развития заболевания будет достаточно мала. По этой причине данную ситуацию можно обозначить как «маркер - классификатор». Типичным примером маркера-классификатора является маммография.

Напротив, в случае редких маркеров (pM << pD) нет смысла проводить массовый скрининг - результаты будут заведомо «нулевые» (SE = pM/pD <<1). Однако высокая диагностическая ценность теста при OR >> 1 позволяет его использовать в клинической практике при наличии дополнительных симптомов и показаний, например в случае неблагоприятной родословной. Подобный маркер можно назвать «маркером - диагностом». Типичный пример - редкие мутации в генах BRCA1 и BRCA1.

Отметим также несколько полезных неравенств, справедливых при любых соотношениях между OR, pM и pD:

SE+SP-1 < (√OR - 1)/(√OR + 1);

PPV+NPV-1 < (√OR - 1)/(√OR + 1); PPV < pD OR.

Работа выполнена при финансовой поддержке Минобрнауки России, ГК №16.612.11.2061 от 15.02.2011.

Литература

Kraft P., Wacholder S., Cornelis M.C. et al., 2009. Beyond odds ratios - communicating disease risk based on genetic profiles // Nature Reviews Genetics. Vol. 10. P. 264-269.

Kraemer H.C., Frank E., Kupfer D.J., 2011. How to assess the clinical impact of treatments on patients, rather than the statistical impact of treatments on measures // International Journal of Methods Psychiatric Research. Vol. 20. P. 63–72.

Jakobsdottir J., Gorin M.B., Conley Y.P., Ferrell R.E., Weeks D.E., 2009. Interpretation of genetic association studies: markers with replicated highly significant odds ratios may be poor classifiers // PLoS Genet. Vol. 5. No 2. e1000337.

Pepe M.S., Gu J.W., Morris D.E., 2010. The potential of genes and other markers to inform about risk // Cancer Epidemiology, Biomarkers & Prevention. Vol. 19. P. 655-665.

Poste G., 2011. Bring on the biomarkers // Nature. Vol. 469. P. 156-157.