

Model Diagnostik dan Prognostik di Bidang Kesehatan Kerja

Eva Suarthana

Centre de Recherche, Hôpital du Sacré-Coeur de Montréal-Université de Montréal

Abstrak

Model prediksi sudah lama digunakan untuk membantu pengambilan keputusan klinis di bidang kedokteran, tetapi relatif baru di bidang kesehatan kerja. Model prediksi dibuat untuk memperkirakan probabilitas suatu kondisi (model diagnostik) atau akan terjadinya suatu kondisi (model prognostik). Model diagnostik dan prognostik dapat dimanfaatkan untuk meningkatkan efisiensi program surveilans pekerja karena model-model tersebut memungkinkan identifikasi kelompok pekerja yang berisiko mengalami penyakit akibat kerja. Dalam tinjauan pustaka ini dibahas hal-hal yang perlu diperhatikan dalam pembuatan model, bagaimana melakukan evaluasi akurasi dan validasi model, transformasi model untuk aplikasi klinis, serta hal-hal yang perlu dicermati dalam implementasi model dalam praktik kesehatan kerja.

Kata kunci: akurasi, kedokteran okupasi, model prediksi, pengambilan keputusan, validitas

Diagnostic and Prognostic Models in Occupational Health

Abstract

Prediction models have long been used to aid clinical decision making, but is relatively new in the field of occupational health. Prediction models are made to estimate the probability of the presence of a condition (diagnostic models) or the occurrence of a condition (prognostic models). Diagnostic and prognostic models can be utilized to improve the efficiency of surveillance programs for workers because these models allows identification of workers at risk of having or developing occupational disease. This review covers aspects that need to be considered in developing a model; evaluating the accuracy and validity of a model; transforming the model for clinical applications; as well as issues in the implementation of the model in occupational health practice.

Key words: accuracy, decision making, occupational medicine, prediction model, validity

Pendahuluan

Model prediksi sudah lama digunakan untuk membantu pengambilan keputusan klinis di kedokteran,^{1,2} tetapi relatif baru di bidang kesehatan kerja. Model prediksi dibuat untuk memperkirakan probabilitas suatu kondisi (model diagnostik) atau akan terjadinya suatu kondisi (model prognostik). Salah satu contoh model diagnostik klinis yang terkenal adalah model yang dibuat oleh Wells untuk memprediksi trombosis vena dalam atau *deep vein thrombosis* (DVT). Wells menunjukkan bahwa informasi yang didapat dari anamnesis, pemeriksaan fisik serta *impedance plethysmography* dapat dengan aman digunakan untuk menyingkirkan DVT. Pemanfaatan model ini mengurangi beban pasien dan biaya perawatan kesehatan karena pasien tidak perlu menjalani venografi yang invasif dan mahal.¹ Skor Framingham merupakan contoh model prognostik yang terkenal.² Skor Framingham dibuat untuk memprediksi terjadinya penyakit jantung koroner (PJK) dalam kurun waktu 10 tahun. Model ini memungkinkan dokter mengidentifikasi pasien dengan probabilitas PJK yang tinggi, sehingga tindakan pencegahan dapat dilakukan sedini mungkin.

Di bidang kesehatan kerja, model diagnostik dan prognostik dapat dimanfaatkan untuk meningkatkan efisiensi program surveilans pekerja. Model prediksi memungkinkan identifikasi kelompok pekerja yang berisiko mengalami penyakit akibat kerja. Beberapa model telah dibuat untuk memprediksi sensitisasi terhadap alergen berat molekul tinggi pada pekerja laboratorium dan pabrik roti,³⁻⁵ serta pneumokoniosis pada pekerja konstruksi⁶. Model regresi logistik diubah menjadi sistem skoring yang mudah dipakai oleh praktisi. Penggunaan model tersebut memungkinkan alokasi sumber daya untuk uji diagnostik lebih lanjut hanya untuk kelompok pekerja

berisiko tinggi.

Sebagai contoh, dalam surveilans pada umumnya, pekerja yang terpajan kristal silika harus menjalani serangkaian uji diagnostik, termasuk kuesioner, fungsi paru, dan rontgen dada. Dengan rendahnya angka prevalensi pneumokoniosis pada kelompok pekerja ini, bisa diprediksi bahwa sejumlah besar pekerja memiliki hasil rontgen dada yang normal. Dengan menggunakan model diagnostik untuk pneumokoniosis, dokter okupasi dapat menggunakan kuesioner dan fungsi paru untuk mengidentifikasi pekerja yang memiliki probabilitas rendah. Kelompok ini tidak memerlukan diagnostik lebih lanjut, akan tetapi dapat diikutsertakan dalam program surveilans selanjutnya.⁶ Rontgen dada dan uji diagnostik lebih lanjut dapat dialokasikan pada pekerja yang memiliki probabilitas yang tinggi. Dengan demikian efisiensi surveilans dapat ditingkatkan.

Pembuatan Model

Hal pertama yang harus diperhatikan dalam membuat suatu model prediksi adalah memilih *outcome* yang relevan dan cara mendefinisikannya.⁷ Sebagai contoh, identifikasi alergi akibat pajanan di tempat kerja tidak harus berfokus pada penyakit alergi misalnya asma akibat kerja, melainkan pada kondisi awal yang terkait dengan asma akibat kerja. Sensitisasi terhadap alergen di tempat kerja merupakan awal terjadinya asma akibat kerja dan dapat diuji dengan mudah.⁸ ⁹ Diagnostik model untuk memprediksi sensitisasi terhadap alergen berat molekul tinggi pada pekerja laboratorium yang terpajan urin tikus dan mencit dapat digunakan untuk mengidentifikasi pekerja yang tersensitisasi, sehingga uji diagnostik lebih lanjut dapat difokuskan pada pekerja dengan probabilitas sensitisasi yang tinggi.⁵

Aspek penting kedua dalam pembuatan model adalah pemilihan prediktor yang

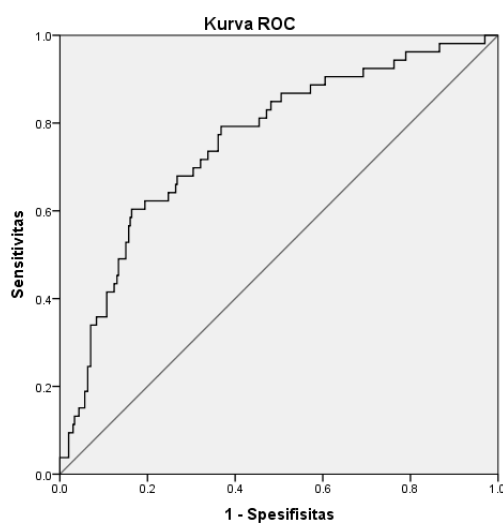
tepat. Berbeda dengan studi etiologi, studi prediksi tidak bertujuan untuk mengidentifikasi hubungan sebab-akibat antara prediktor dan *outcome*.¹⁰ Selain itu, urutan waktu (penyebab mendahului akibat) tidak diperlukan dalam studi diagnostik. Secara inheren studi diagnostik adalah studi potong lintang dan tidak jarang prediktor yang digunakan dalam model diagnostik merupakan konsekuensi dari suatu penyakit. Sebagai contoh, kebiasaan merokok terpilih sebagai prediktor dalam model diagnostik untuk pneumokoniosis. Kebiasaan merokok memiliki rasio odds (OR) sebesar 2,4. Hal itu berarti bahwa, dibandingkan dengan non-perokok, mereka memiliki 2,4 kali lebih tinggi dari probabilitas orang yang memiliki rontgen dada yang mengindikasikan adanya pneumokoniosis, tanpa mengacu pada hubungan kausalitas antara merokok dan pneumokoniosis.⁶

Dalam praktiknya, diagnosis jarang sekali ditegakkan berdasarkan hasil tes tunggal. Di lain pihak, banyak tes menghasilkan informasi yang serupa. Oleh karena itu, analisis multivariabel regresi digunakan untuk mengevaluasi nilai (*added value*) diagnostik atau prognostik suatu tes dibandingkan tes lain.¹⁰ Untuk menentukan prediktor pada uji multivariabel, peneliti bisa

melakukan kajian literatur. Peneliti juga bisa memasukkan prediktor dengan nilai $P < 0,25$ pada uji univariabel. Nilai P yang kecil ($< 0,05$) tidak dianjurkan untuk menghindari eliminasi prediktor yang penting dalam seleksi awal ini. Model multivariabel final dapat dipilih dengan metode *backward stepwise selection* menggunakan $P < 0,157$ (*Akaike criterion*) untuk inklusi dan eksklusi prediktor final.¹⁰

Akurasi dan Validasi Model

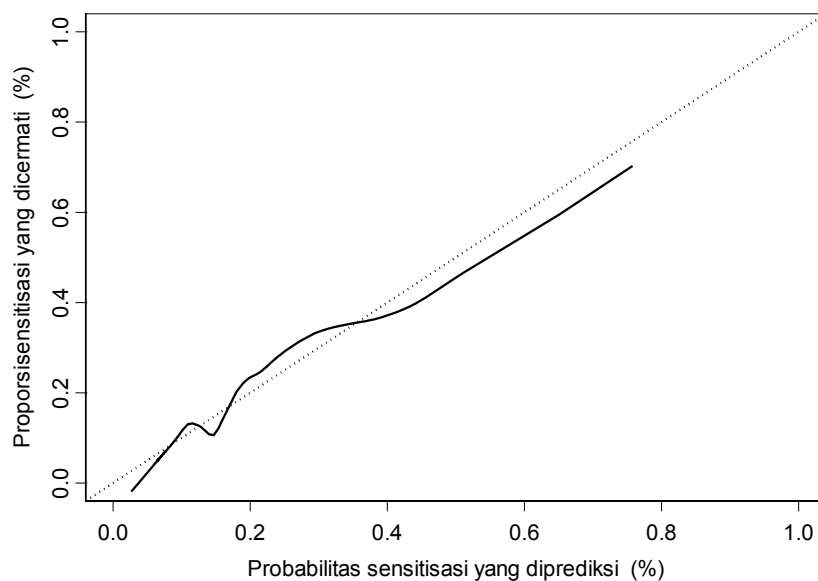
Akurasi model final dinilai menggunakan parameter diskriminasi dan kalibrasi. Diskriminasi adalah kemampuan suatu model untuk membedakan individu dengan dan tanpa kondisi yang diteliti. Kemampuan ini dievaluasi menggunakan kurva *receiver operating characteristics* (ROC). Kurva ROC menghubungkan *true positive* (sensitivitas) dan *false positive* (1- spesifisitas) berbagai titik potong probabilitas yang dihasilkan oleh suatu model. Daerah di bawah kurva (AUC) berkisar mulai 0,5, yang berarti model tidak memiliki kemampuan diskriminasi, hingga 1,0 (diskriminasi yang sempurna).¹¹ Contoh kurva ROC ditampilkan pada Gambar 1.



Gambar 1. Kurva ROC model diagnostik untuk memprediksi sensitivitas terhadap alergen hewan laboratorium. Daerah di bawah garis diagonal (garis referensi) luasnya 0,5. Daerah di bawah kurva ROC (AUC) model ini adalah 0,757 (95% CI: 0,686-0,828) menunjukkan kemampuan diskriminasi yang cukup baik.

Kalibrasi adalah kesesuaian antara probabilitas yang diprediksi dengan frekuensi luaran (*outcome*) yang diamati. Kalibrasi dievaluasi secara grafis dengan menghubungkan probabilitas diprediksi (aksis X) dengan proporsi luaran yang diamati (aksis Y). Dengan menggunakan perangkat lunak S-Plus atau R, peneliti bisa mendapatkan garis kalibrasi. Kalibrasi yang sempurna dilambangkan dengan garis yang memiliki $slope=1$, $intercept=0$, dan melintang 45° memotong sumbu $Y=0$.¹² Contoh garis kalibrasi ditampilkan pada Gambar 2. Kalibrasi juga dapat dievaluasi menggunakan uji *goodness of fit* Hosmer-Lemeshow. Nilai $P > 0,05$ menunjukkan tidak ada perbedaan bermakna antara probabilitas yang diprediksi dengan frekuensi *outcome* yang diamati.¹³

luas, validasi eksternal perlu dilakukan pada populasi baru yang berasal dari domain yang sama dengan populasi yang digunakan untuk membuat model. Contohnya, sebuah model diagnostik untuk sensitisasi terhadap alergen hewan laboratorium dibuat pada populasi pekerja laboratorium di Belanda. Model ini divalidasi pada peserta sekolah kejuruan di Kanada yang juga terpajan alergen hewan laboratorium.¹⁵ Ternyata kemampuan diskriminasi model di populasi pelajar Kanada seimbang dengan populasi pekerja Belanda. Akan tetapi, kalibrasi di populasi pelajar Kanada ini kurang baik karena perbedaan prevalensi luaran antara kedua populasi. Setelah dilakukan modifikasi *intercept*, model menjadi akurat dan sah di populasi pelajar Kanada.



Gambar 2. Grafik kalibrasi model diagnostik untuk memprediksi sensitisasi terhadap alergen hewan laboratorium. Garis kalibrasi (hitam tebal) memiliki $intercept=0.16$ dan $slope=1.09$ dan melintang mendekati garis diagonal menunjukkan kalibrasi yang baik.

Secara umum, model diagnostik atau prognostik menunjukkan akurasi yang lebih rendah pada populasi baru.¹⁴ Oleh karena itu, sebelum suatu model dapat digunakan secara

Dalam kenyataan, tidak setiap saat kita bisa mendapatkan data untuk melakukan validasi eksternal. Pada kondisi seperti itu, setidaknya peneliti melakukan validasi

internal dengan menggunakan teknik *bootstrapping*. Dari teknik itu akan dihasilkan faktor koreksi yang nilainya berkisar antara 0 dan 1. Koefisien regresi dan *intercept* model multivariabel final dimultiplikasi dengan faktor koreksi ini untuk mencegah terjadinya *over optimism* (prediksi yang terlalu tinggi atau terlalu rendah).¹⁰ Model yang memiliki faktor koreksi mendekati angka satu dikatakan memiliki validitas internal yang baik. Dari *bootstrapping* juga akan didapatkan nilai AUC yang dikoreksi (*c-statistic*). *Bootstrapping* dapat dilakukan dengan perangkat lunak S-Plus maupun STATA.

Aplikasi model

Setelah kita memiliki model yang valid dan akurat, kita bisa mengubahnya menjadi sistem skor atau nomogram agar mudah digunakan oleh klinisi dalam praktek sehari-hari. Jika semua prediktor bersifat kategorikal, model dapat diubah menjadi sistem skor. Jika prediktor ada yang bersifat kontinyu, model bisa ditransformasi menjadi nomogram. Untuk membuat sistem skor, koefisien regresi setiap prediktor di model final dibagi dengan koefisien regresi terkecil di model tersebut, lalu dibulatkan.^{6,8,9} Nomogram bisa didapat dari perangkat lunak S-Plus. Contoh nomogram ditampilkan pada Gambar 3.

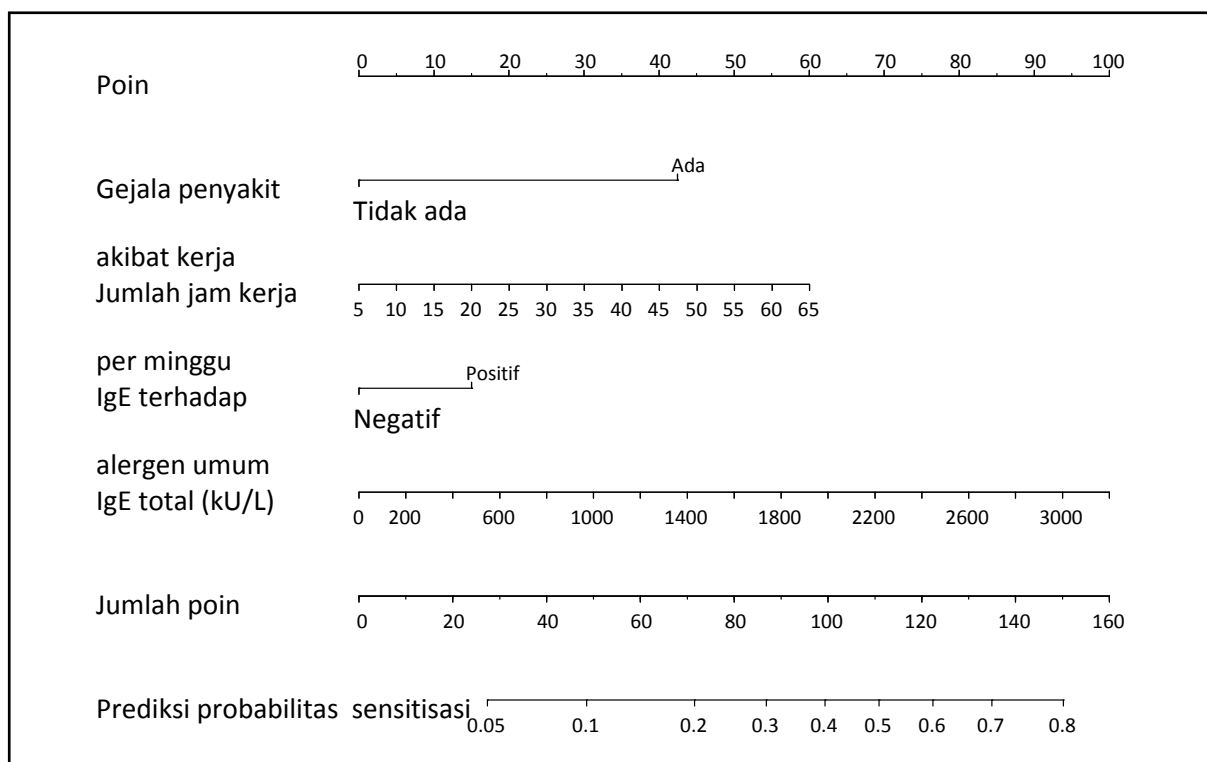
Langkah berikutnya adalah untuk menentukan ambang batas (*cut-off*) probabilitas untuk membuat kelompok risiko (tinggi-rendah, tinggi-sedang-rendah). Nilai ambang ini akan mempengaruhi kesalahan klasifikasi *outcome*. Pada umumnya, makin tinggi nilai ambang yang kita pilih, maka makin rendah persentase subjek yang termasuk ke dalam kelompok risiko tinggi (dengan kata lain jumlah rujukan lebih rendah); spesifisitas makin tinggi (angka positif palsu rendah); sedangkan sensitivitas makin rendah (angka negatif palsu lebih

tinggi). Oleh karena itu, pilihan nilai ambang harus berdasarkan pada keseimbangan antara proporsi kasus yang hilang (negatif palsu) dan jumlah rujukan yang tidak perlu (positif palsu). Peneliti dapat melakukan *decision curve analysis* (DCA), metode yang relatif baru dan sederhana untuk mengevaluasi model prediksi. DCA dapat digunakan untuk membantu dokter membuat keputusan untuk melakukan atau tidak melakukan uji diagnostik yang invasif dan mahal, ataupun untuk memutuskan pemberian terapi dengan cara mengevaluasi konsekuensi (*relative harm*) positif palsu dibandingkan negatif palsu.¹⁶

Prospek Model Prediksi dalam Praktik Kesehatan Kerja

Penerapan model prediksi dalam praktik kesehatan kerja adalah pendekatan baru. Deteksi awal pekerja yang mengalami penyakit atau kondisi subklinis akibat kerja dapat mengurangi beban penyakit akibat kerja. Yang penting, peneliti harus memperhatikan aspek etika dan hukum dalam merancang model dan menetapkan konteks di mana dan bagaimana model tersebut diaplikasikan. Model prediksi jangan disalahgunakan untuk *pre-employment screening*, misalnya untuk menolak calon pekerja yang berisiko tinggi mengalami penyakit akibat kerja. Penggunaan model harus dibatasi untuk program surveilans pekerja, sehingga dokter okupasi bisa mengidentifikasi pekerja yang berisiko tinggi dan melakukan intervensi sedini mungkin.

Kesempatan terbuka lebar untuk mengembangkan model prediksi pada penyakit akibat kerja yang beragam, terutama mengingat banyaknya studi epidemiologi berskala besar. Penelitian selanjutnya juga diperlukan untuk mengevaluasi efektivitas penerapan model ini dalam praktik kesehatan kerja.



Gambar 3. Nomogram model diagnostik untuk memprediksi sensitisasi terhadap alergen hewan laboratorium. Garis paling atas menunjukkan poin yang didapat untuk setiap prediktor dalam model. Misalnya seorang pekerja yang mengalami sesak napas di tempat kerja, bekerja sebanyak 35 jam per minggu, memiliki IgE positif terhadap polen, dan nilai IgE total 300 kU/L akan mendapatkan poin $42+30+15+13$. Jumlah poin adalah 100 dengan demikian prediksi probabilitas sensitisasi pekerja ini adalah 0.4 (40%).

Daftar Pustaka

1. Wells PS, Anderson DR, Bormanis J, Guy F, Mitchell M, Gray L, et al. Value of assessment of pretest probability of deep-vein thrombosis in clinical management. *Lancet*. 1997;350(9094):1795-8.
2. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. *Circulation*. 1998;97(18):1837-47.
3. Suarathana E, Meijer E, Heederik D, Ghezzi H, Malo JL, Gautrin D. The Dutch diagnostic model for laboratory animal allergen sensitization was generalizable in Canadian apprentices. *J Clin Epidemiol*. 2009;62(5):542-9.
4. Suarathana E, Vergouwe Y, Moons KG, de Monchy J, Grobbee D, Heederik D, et al. A diagnostic model for the detection of sensitization to wheat allergens was developed and validated in bakery workers. *J Clin Epidemiol*. 2010;63(9):1011-9.
5. Suarathana E, Vergouwe Y, Nieuwenhuijsen M, Heederik D, Grobbee DE, Meijer E. Diagnostic model for sensitization in workers exposed to occupational high molecular weight allergens. *Am J Ind Med*. 2005;48(3):168-74.
6. Suarathana E, Moons KG, Heederik D, Meijer E. A simple diagnostic model for ruling out pneumoconiosis among construction workers. *Occup Environ Med*. 2007;64(9):595-601.
7. Coggon D MC, Palmer KT, Evanoff B. Assessing case definitions in the absence of a diagnostic gold standard. *Int J Epidemiol*. 2005;34:949-52.
8. Meijer E, Grobbee DE, Heederik D. A strategy for health surveillance in laboratory animal workers exposed to high molecular weight allergens. *Occup Environ Med*. 2004;61(10):831-7.
9. Gautrin D, Ghezzi H, Infante-Rivard C, Magnan M, L'Archeveque J, Suarathana E, et al. Long-term

- outcomes in a prospective cohort of apprentices exposed to high-molecular-weight agents. *Am J Respir Crit Care Med.* 2008;177(8):871-9.
10. Harrell FE, Jr., Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15(4):361-87.
 11. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology.* 1982;143(1):29-36.
 12. Miller M, Langefeld C, Tierney W. Validation of probabilistic predictions. *Med Decis Making.* 1993;13:49-58.
 13. Hosmer D, Lemeshow S. *Applied logistic regression.* New York: John Wiley and Sons, Inc, 1989.
 14. Justice AC, Covinsky KE, Berlin JA. Assessing the generalizability of prognostic information. *Ann Intern Med.* 1999;130(6):515-24.
 15. Suarathana E, Meijer E, Heederik D, Ghezzi H, Malo JL, Gautrin D. The Dutch diagnostic model for laboratory animal allergen sensitization was generalizable in Canadian apprentices. *J Clin Epidemiol.* 2008.
 16. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making.* 2006;26(6):565-74.