Autoantibodies as predictors of disease

David Leslie,1 Peter Lipsky,2 and Abner Louis Notkins3

1Department of Diabetes and Metabolism, St. Bartholomew’s Hospital, London, United Kingdom
2Laboratory of Autoimmunity, National Institute of Arthritis and Musculoskeletal and Skin Diseases, and
3Experimental Medicine Section, Oral Infection and Immunity Branch, National Institute of Dental and Craniofacial Research,
NIH, Bethesda, Maryland, USA

Address correspondence to: David Leslie, Department of Diabetes and Metabolism, St. Bartholomew’s Hospital, London EC1A 7BE, United Kingdom. Phone: 44-0-207-601-7450; Fax: 44-0-207-601-7449; E-mail: r.d.g.leslie@mds.qmw.ac.uk.

Activation of the immune response is a major feature of many disease processes. Immune responses can be protective, as in infectious diseases, or destructive, as in autoimmune inflammatory diseases, or both. The immune response usually involves activation of both T and B cells, the latter producing antibodies that can be detected in the sera and can be used to guide the clinical management of certain diseases. Here, we focus on autoantibodies as predictive markers of disease. While the practical value of autoantibodies has been realized in some clinical conditions, it remains underutilized in the majority of diseases. Recognizing the clinical potential of autoantibodies and identifying appropriate populations to screen for such autoantibodies, we argue, could have rich practical rewards.

Autoantibodies as markers of disease activity and severity

Antibodies may reflect the presence, nature, and intensity of the immune response. Since in autoimmune diseases the immune response is itself part of the disease process, it is possible to use autoantibodies as markers of disease activity. Autoantibodies can be detected in diseases with a long prodrome during which there are no clinical symptoms. In some of these diseases autoantibodies can predict both the likelihood of clinical disease and the rate of progression to disease. Hence, many such autoimmune diseases are long-term and chronic in nature (1, 2). A variety of studies have now shown a direct relationship between the titer of the autoantibodies and the severity of the disease (1–3).

Autoantibodies as markers to define and classify disease

In patients with established disease, autoantibodies can help define the nature of the disease and provide markers to classify the disease. For example, type I diabetes, thyroiditis, and adrenalitis are classified as autoimmune or not autoimmune, based on the presence or absence of disease-associated antibodies. Similarly, there are several causes of atrophic gastritis and of vitamin B12 deficiency, but the combination of the two, in association with autoantibodies to parietal cells or intrinsic factor, indicates that the cause is autoimmune gastritis, also called pernicious anemia (PA) (4).

Autoantibodies as markers to predict disease

Since autoantibodies are markers of disease activity, it follows that, at least under some circumstances, autoantibodies should be able to predict disease. This approach is especially promising for diseases with a long preclinical period, a feature of many organ-specific autoimmune diseases. The aim of disease prediction is disease prevention. Autoimmune diseases, which affect at least 5% of the population, might be prevented by avoiding the environmental factors that trigger the disease (primary prevention) or by use of therapy that modulates the destructive process before the onset of clinical symptoms (secondary prevention). Accurate disease prediction is vital for secondary prevention, so that therapy is given only to those individuals who are likely to become clinically ill.

Three parameters must be carefully quantitated for predictive tests to be clinically useful: sensitivity of prediction, specificity of prediction, and positive predictive values.

Sensitivity of prediction is calculated by dividing the number of subjects in a cohort with autoantibodies who develop a disease by the overall number of subjects who develop the disease. Ideally, every subject who develops an autoimmune disease will have that particular autoantibody (high disease sensitivity) before clinical diagnosis. However, since disease-associated autoantibodies do not develop simultaneously and since many patients have only one antigen-specific autoantibody, using a panel of different autoantibodies is likely to increase the sensitivity of prediction.
Specificity of prediction with an autoantibody marker reflects the chance that a person without that marker will remain disease-free. It is calculated by dividing the number of subjects in a cohort without that autoantibody marker who do not go on to develop disease by the total number of subjects who do not develop the disease. Specificity is important if a disease marker is to be used to identify individuals either for counseling or for therapy to prevent the disease from developing. A reciprocal relationship exists between sensitivity and specificity (5). The higher the threshold for autoantibody positivity based on the normal population, the more specifically the autoantibody assay identifies patients with clinical disease, but at the cost of excluding many patients with low autoantibody signals.

If an autoantibody is to be used to predict disease, then ideally every subject with the autoantibody, but without clinical disease, will eventually develop clinical disease. That is, the test should display high disease positive predictive value. The positive predictive value is calculated by dividing the number of autoantibody-positive subjects in the initial sample who go on to develop clinical disease by the overall number of autoantibody-positive subjects. The prognostic significance of any marker varies in populations at differing levels of risk. If the disease risk is high then the predictive power can be high, but when the disease risk is low, as in the general population, then there is a corresponding reduction in predictive power (5). Predictions based on cross-sectional analyses must be verified in prospective studies. Many studies use predictive values based on cross-sectional data of cases with established clinical disease. In general, this approach is invalid and relates to identification of disease cases, not prediction.

Prediction of specific autoimmune diseases
The idea that autoantibodies can be used to predict disease comes from extensive studies on type 1 diabetes and to a lesser degree from studies on other diseases. The information below on several representative diseases illustrates the current state of the art and the promise and problems associated with autoantibody prediction.

Type 1 diabetes
Glutamic acid decarboxylase (GAD), the protein tyrosine phosphatase–like molecule, IA-2, and insulin are the principal autoantigens in type 1 diabetes (1). As discussed by Notkins and Lernmark in this Perspective series (6), these autoantibodies have been used to study disease activity, determine the rate of disease progression, and help classify and predict clinical disease. The appearance of autoantibodies in young children led to the proposal that type 1 diabetes is induced in early life (7). Studies of autoantibodies during the prediabetic period showed variability in the rate of progression to clinical diabetes based on the titer and number of antibodies. For example, of 3,578 first-degree relatives of patients with type 1 diabetes, 105 subsequently developed diabetes and all but seven had antibodies to IA-2, GAD, or insulin years before the development of the disease, giving a sensitivity of prediction of 92% (8–11). Thus far, all prospective studies on relatives of type 1 diabetes patients have shown that the combination of two or more autoantibodies gives a higher positive predictive value than any single autoantibody. The positive predictive values in relatives with one, two, or three autoantibodies were 2%, 25%, and 70% respectively in one study and 15%, 44%, and 100% respectively in another study (8, 9). Moreover, diabetes is now being classified as autoimmune or not autoimmune, largely based on the presence or absence of autoantibodies. Up to 90% of initially non–insulin-requiring, but autoantibody-positive, diabetes patients have progressed to insulin dependence within 6 years (12). Autoantibody screening is now being widely used to admit potential diabetic subjects into therapeutic intervention trials.

Thyroiditis
Thyroid autoantibodies can also reflect disease activity and progression and be valuable in disease prediction and classification (see Rapoport and McLachlan, this Perspective series, ref. 13). There are two major clinical diseases associated with thyroid autoimmunity, hypothyroidism (Hashimoto thyroiditis) and hyperthyroidism (Graves disease). Two of the principal thyroid autoantigens in the former are thyroid peroxidase and thyroglobulin. Autoantibodies to these antigens rarely develop before 20 years of age, but they may presage subsequent clinical disease (primary hypothyroidism) (2). The probability of developing overt hypothyroidism within the next 20 years in women who are thyroid peroxidase antibody–negative with a thyroid stimulating hormone (TSH) less than 2 mU/l is less than 5% (14). That probability increases logarithmically to 55% when the TSH is greater than 6 mU/l in thyroid peroxidase antibody–positive subjects, an annual rate of progression of 4.3%, as compared with just 2.6% for those with only elevated TSH and 2.1% for those with only peroxidase autoantibodies (14). Higher rates of progression have been found in men, who are five times more likely than women to progress to overt disease; in women aged 45 years or older; in patients with TSH levels greater than 20 mU/l; and in patients with thyroid antibody titers greater than 1:100,000 (14, 15). About 10% of thyroid antibody–positive subjects in these studies subsequently became antibody-negative, and in some subjects mildly elevated TSH returned to normal. These observations in thyroiditis mimic some of the features associated with diabetes autoimmunity, including the dual parameter model in which evidence of both autoimmunity and target organ failure are predictive of progression to clinical disease and the risk of progression to disease is related to antibody titer and persistence. These findings are consistent with the broad spectrum of clinical consequences of thyroid autoimmunity,
which extends even to hyperthyroidism due, for example, to TSH receptor autoantibodies.

Celiac disease
Celiac disease is an autoimmune disease associated with total or sub-total villous atrophy leading to malabsorption from the gut (16). As with diabetes, antibodies associated with celiac disease can predict progression to clinical disease and can be used to define the cause of malabsorption. The disease is associated with endomysial antibodies and antibodies to gliadin and reticulin (see article by Papadopoulos et al., this Perspective series, ref. 17). Gliadin antibodies, both IgA and IgG, have been used in screening and have a sensitivity and specificity for prediction of celiac disease of 70–100%, but most of the relevant studies have been cross-sectional, testing established cases and controls; thus, these estimates are not true estimates of prediction but of identification (16, 18, 19). Serum reticulin antibodies, notably IgA antibodies against R1-type reticulin, have a sensitivity and specificity at diagnosis in excess of 90% (20). Endomysial antibodies are also highly predictive of celiac disease (18, 19, 21); these antibodies identify tissue transglutaminase (tTG), which uses gliadin as a preferred substrate. The specificity of tTG antibody for celiac disease has been found to be 99.5% with a sensitivity of 95.6% (18). In one study, population screening using IgA endomysial and anti-reticulin antibodies showed a sensitivity of 100% for the antibodies when sera were persistently positive over a 4-year period, but the positive predictive value was only 27% when based on positivity at the initial screening. Interestingly, none of the subjects with transient antibodies had villous atrophy, suggesting that this feature develops only after chronic immune activation (22). In another study of patients referred by general practitioners with nonspecific abdominal symptoms, IgA endomysial antibodies had 100% positive predictive value for celiac disease as confirmed by small bowel biopsy (23). Thus, as with type 1 diabetes and thyroiditis, a substantial proportion have transient autoantibodies, but when the autoantibodies persist, the risk of progression to clinical celiac disease is high. As a result, celiac disease–associated autoantibodies are now widely used for disease prediction and diagnosis. Indeed removal of the antigen, gluten, is currently the therapy of choice for celiac disease.

Adrenalitis
Adrenalitis is another organ-specific autoimmune disease that can lead to primary adrenal gland failure (Addison disease). Adrenalitis-associated antibodies react with 21-hydroxylase, 17-hydroxylase, and the cytochrome P450 side chain cleavage enzyme (3). In at least a proportion of the cases, these antibodies first appear in childhood. As with diabetes, antibodies associated with adrenalitis can predict progression to clinical disease and have been used to classify the cause of primary adrenal failure. The risk of progression to adrenal failure is dependent on antibody titer,

age, and HLA alleles (3, 24, 25). In a prospective study of 808 children without adrenal failure, ten children had 21-hydroxylase antibodies and nine developed adrenal failure within 3 to 121 months. In contrast, none of the antibody-negative children developed adrenal failure (3). Progression to Addison disease in subjects with adrenal antibodies was observed in only 21% of adults (91% of whom had 21-hydroxylase antibodies) but in 90% of children (3, 24). In selected patients with type 1 diabetes or Addison disease, those with both 21-hydroxylase autoantibodies and the DRB1*0404 allele were at higher risk of Addison disease (80%) than those with these autoantibodies but with DRB1*0401 or DRB1*0402 alleles (10%) (25).

Pernicious anemia
Pernicious anemia (PA) is an autoimmune disease leading to blockade in gastric hydrochloric acid secretion and vitamin B12 deficiency (4). Autoantibodies are directed against gastric H+/K+-ATPase and intrinsic factor. Gastric H+/K+-ATPase is responsible for secretion of hydrogen ions in exchange for potassium ions by the membranes of parietal cells. Autoantibodies bind to the 100-kDa catalytic α subunit and the 60- to 90-kDa glycoprotein β subunit of this enzyme. These autoantibodies are detected in about 90% of patients with PA but also in about 13% of their nonanemic first-degree relatives (26). Intrinsic factor, a 60-kDa glycoprotein, binds to and is required for absorption of B12. Two types of autoantibodies bind to intrinsic factor. Of these autoantibodies, type 1 autoantibody is more common, blocks the binding of vitamin B12 to intrinsic factor, and is found in about 70% of PA patients. In the general population, the prevalence of parietal cell autoantibodies, as determined by immunofluorescence, increases with age from about 3% at age 30 to 10% at age 80 (4). In a cross-sectional study of atrophic gastritis in first-degree relatives of PA patients, parietal cell antibodies had a disease identification specificity of 87%, and sensitivity of 65%, but a positive predictive value of only 44% (27). Again it should be noted that these figures relate to identification of disease and not strictly to prediction. Other studies have shown that as with diabetes and thyroid disease, the combination of autoantibodies and functional changes identifies disease with a higher level of precision (28, 29). HLA may also be important; a study of patients with type 1 diabetes revealed parietal cell antibodies in 21% of the patients, and the presence of these antibodies was associated with the HLA DR5 allele (29). Long-term prospective studies are needed to more definitively define the value of these autoantibodies in prediction.

Rheumatic diseases
In rheumatic diseases, autoantibodies have been used most often to confirm diagnosis and on occasion to predict prognosis or organ involvement. Some limited data suggest that they can be used to predict the development of disease. Rheumatic autoimmune diseases
are relatively uncommon conditions, each affecting less than 1% of the population. The most common autoantibodies, namely rheumatoid factor (RF), an autoantibody to the Fc portion of IgG, and anti-nuclear antibodies (ANAs) are frequently found in rheumatoid arthritis (<75%) and systemic lupus erythematosus (SLE) (90–100%), respectively. However, these autoantibodies are also found in 5–10% of normal individuals, especially the elderly. For this reason, the number of normal individuals with these autoantibodies exceeds the number of patients with these autoantibodies by many-fold. This observation has discouraged the use of these particular autoantibodies in screening individuals for the likelihood of developing rheumatic disease. Moreover, uncertainty relating to the role of these autoantibodies in disease pathogenesis has further diminished enthusiasm for their use in screening asymptomatic individuals.

For the past 50 years, the relationship of the production of autoantibodies and the manifestations of rheumatic diseases has been intensively studied. Most of the focus has been on the association with specific diseases. Many of the initial autoantibodies studied, including RF and ANA, are not highly specific for particular rheumatic diseases or even for rheumatic diseases in general, being found in a variety of chronic inflammatory conditions as well as in a small percentage of apparently normal individuals (30, 31). Their major use, therefore, became confirmation of diagnosis in patients who presented with many clinical features of specific rheumatic disease. Nevertheless, the presence of some of these autoantibodies at diagnosis can provide useful prognostic information. For example, patients with rheumatoid arthritis and high titers of RF tend to have a more rapidly progressive course (32, 33). By contrast, a positive ANA test provides no prognostic information, although it is associated with increased likelihood of a diagnosis of SLE in patients who show some clinical features of the disease.

More recently, a variety of autoantibodies have been shown to be highly specific for certain rheumatic diseases, including anti–double-stranded DNA (34), anti-Sm (35), and anti–ribosomal P (36) antibodies in SLE, anti–topoisomerase I (37) in scleroderma, autoantibodies against citrulline-modified proteins in rheumatoid arthritis (38), and anti–tRNA synthetase antibodies in myositis (39). Although many of these autoantibodies are highly specific for a particular rheumatic disease, their sensitivity as diagnostic markers tends to be lower. Their value as screening tests in asymptomatic individuals has not been thoroughly evaluated, although they might be more useful in this regard than less specific tests. Although these autoantibodies are commonly used solely for diagnosis and/or prognosis, a few reports have suggested that they may also be predictive of rheumatic disease. In general, however, these relationships have not been tested in large-scale population studies. For example, the presence of RF in asymptomatic individuals may predict the subsequent development of rheumatoid arthritis (40–42). One of these studies (42), involving a large cohort of Pima Indians assessed over nearly 20 years, found a correlation between the RF titer and development of RA. Similarly, anti–topoisomerase I autoantibodies may precede the development of scleroderma. Furthermore, the development of myositis-specific antibodies may be predictive of the development of polymyositis (39, 43, 44). Finally, the best documented relationship between the presence of specific autoantibodies and the subsequent development of an inflammatory or autoimmune disease with rheumatic features is the relationship between the presence of autoantibodies to the E2 component of pyruvate dehydrogenase and primary biliary cirrhosis (PBC) (45, 46). The presence of this autoantibody can precede the development of PBC by many years, and many asymptomatic individuals with these autoantibodies subsequently develop PBC. Although only demonstrated in a small number of examples, the results suggest that the presence of certain autoantibodies may have some predictive utility in rheumatic disease. In general, the relationships have not been explored sufficiently to be totally convincing, but preliminary results suggest that additional investigation is warranted.

Concluding comments

A number of broad principles that may apply to autoimmune diseases in general can be drawn from the observations on these selected autoimmune diseases. First, autoantibodies reflect the disease process and, in those cases where the disease has a long prodrome, antibodies can predict clinical disease. This is particularly relevant when the autoantibody itself has the capacity to damage tissue. Second, a number of distinct autoantibodies are associated with any one disease, and some are more predictive of progression to clinical symptoms than others. Third, the risk of progression to disease, the rate of progression, and the severity of the clinical disease can be predicted to a degree by the number of autoantibodies, the type of autoantibody, the titer of the autoantibody, and associated features including genetic risk and evidence of target organ failure.

These observations hold out the prospect of screening the general population to identify individuals at high risk for some autoimmune diseases. Such cases might be amenable to therapy either to prevent progression to clinical disease or to limit the impact of disease. However, a number of questions must be answered before screening becomes a reality and for appropriate strategies to be devised. For example, the predictive values of autoantibodies will likely be different when assessed in twins, in first-degree relatives, or in the general population. In addition, the best age at which to screen for disease-associated autoantibodies varies with the particular disease. For example, many of the diabetes-associated autoantibodies appear by 5 years of age (1), so screening should ideally be performed at birth and repeated at intervals thereafter. Limiting the screening program to only young children, however, would miss a significant fraction of
potential cases. Thyroid antibodies, on the other hand, rarely appear before 20 years of age and there would be no value in screening for them until after that age. A number of other diseases not discussed here are associated with the presence of autoantibodies, but there are no large screening studies to determine whether these same autoantibodies can predict the disease; for example, neurological disorders such as myasthenia gravis and Lambert-Eaton syndrome, in which autoantibodies can be diagnostic of the disease. Nor is information available on the ideal time to screen for autoantibodies in rheumatic diseases or whether the approach is feasible and rational. Large population-based screening studies have been shown to be feasible for diabetes-associated autoantibodies, so similar studies using a wider panel of autoantibodies associated with other diseases may prove valuable. The lack of long-term prospective studies is perhaps the single most important reason why the role of autoantibodies as predictors of disease is still in its infancy.

By identifying at birth those infants who are genetically at increased or reduced risk of autoimmune diseases, it might be possible to substantially reduce the numbers in a population who would have to be screened for autoantibodies. This additional step might also increase the predictive value of a positive autoantibody test. For example, non-diabetic relatives with autoantibodies are unlikely to develop diabetes if they have the HLA allele DQB1*0602 (7), whereas individuals with 21-hydroxylase autoantibodies and the HLA DRBI*0404 allele are at particularly high risk of progression to adrenal failure. Performance characteristics of autoantibody assays also might have an important influence on their predictive value, so improving assay characteristics could improve their predictive value. Maximal predictive sensitivity and specificity may require testing of different sets of autoantibodies at different ages and at different times in the course of a particular disease. Novel antigen-specific autoantibodies, once identified, might improve prediction even further. Autoantibodies also might be used as surrogate markers to follow disease progression, and their disappearance upon therapy might indicate a beneficial response.

In conclusion, although many issues remain unresolved, screening of populations for susceptibility to certain autoimmune diseases is now feasible. Apart from the studies on type 1 diabetes, however, the value of autoantibodies as predictors of disease has not been fully explored, nor has the potential been fully realized. It is likely that in the future, risk assessment will use mathematical models incorporating the number and character of autoantibodies together with genetic markers. High throughput methods should make it possible to rapidly screen for dozens of autoantibodies at low cost, and screening for autoantibodies may become a routine part of a medical examination. The practical value of screening healthy populations to detect individuals at high risk for a particular autoimmune disease will be enhanced enormously once preventive measures and safe therapy become available.


