CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Latent Tuberculosis Infection Testing in People with Compromised Immunity Prior to Biologic Therapy: A Review of Diagnostic Accuracy, Clinical Utility, and Guidelines

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	December 1, 2020
Report Length:	47 Pages

CADTH

Authors: Calvin Young, Melissa Severn

Cite As: Latent Tuberculosis Infection Testing in People with Compromised Immunity Prior to Biologic Therapy: A Review of Diagnostic Accuracy, Clinical Utility, and Guidelines. Ottawa: CADTH; 2020 Nov. (CADTH rapid response report: summary with critical appraisal).

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Funding: CADTH receives funding from Canada's federal, provincial, and territorial governments, with the exception of Quebec.

Questions or requests for information about this report can be directed to **Requests@CADTH.ca**

Abbreviations

AGREE	Appraisal of Guidelines for Research and Evaluation
AMSTAR 2	A MeaSurement Tool to Assess systematic Reviews 2
CI	confidence interval
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IGRA	interferon gamma release assay
NICE	National Institute for Health and Care Excellence
SE	standard error
SIGN	Scottish Intercollegiate Guidelines Network
TB	tuberculosis
TNF	tumour necrosis factor
TST	tuberculin skin test
QUADAS-2	Quality Assessment of Diagnostic Accuracy Studies 2

Context and Policy Issues

Tuberculosis (TB) is a chronic communicable disease that primarily affects the respiratory system. It is caused by infection with Mycobacterium tuberculosis, a species of bacteria in the Mycobacteriaceae family. As much as 25% of the global population is estimated to have TB infection,[1] although only a small portion of infected individuals are ill as a result of the infection. The incidence of TB varies greatly by country. For example, the rate of active TB in Canada is approximately 4.9 per 100,000 population, although this rate is considerably higher in Canadian-born Indigenous populations (21.5 per 100,000 population).[2] By comparison, the incidence of active TB in the Philippines, which is classified as a high-burden country, is estimated to be 554 per 100,000 population.[3]

There are two stages to TB: 1) latent TB infection, and 2) active TB infection. Latent TB typically occurs immediately following exposure to the bacteria. Although an individual who has latent TB infection will not exhibit any symptoms and they are not considered contagious, there are dormant bacteria residing in their lungs.[4] If latent TB is left untreated it can progress to active infection. Those who have active infection are contagious and may show symptoms such as coughing, chest pain, unexpected weight loss, loss of appetite, night sweats, fever, fatigue, and chills.[4] Without medical intervention (e.g., rifampin, isoniazid, pyrazinamide, ethambutol), TB can be fatal.[5] Individuals with who have HIV, are transplant recipients, are on biologic therapies (e.g., TNF [tumour necrosis factor] inhibitors), and those who have impaired or compromised immune systems for other reasons are at an increased risk for latent TB reactivation.[6]

There is no established gold standard technique for diagnosing latent TB infection.[7] The most commonly used screening methods are the tuberculin skin test (TST) and interferon gamma release assays (IGRAs); however, these diagnostic tools have limited sensitivity and specificity.[8] The objective of the current report is to evaluate the evidence regarding the

diagnostic accuracy and clinical utility of the TST for patients with compromised immunity prior to initiating biologic treatment. Additionally, evidence-based guidelines regarding testing for latent TB infection in patients with compromised immunity prior to initiating biologic therapy will be reviewed. This report expands upon a previously completed CADTH report (list of references).[9]

This report is a component of a larger CADTH Condition Level Review on TB. A condition level review is an assessment that incorporates all aspects of a condition, from prevention, detection, treatment, and management. For more information on CADTH's Condition Level Review of TB, please visit the project page (https://www.cadth.ca/tuberculosis).

Research Questions

1. What is the diagnostic accuracy of using the tuberculin skin test for patients with compromised immunity prior to initiating biologic treatment?
2. What is the clinical utility of using the tuberculin skin test for patients with compromised immunity prior to initiating biologic treatment?
3. What are the evidence-based guidelines regarding the testing for latent tuberculosis infection in patients with compromised immunity prior to initiating biologic therapy?

Key Findings

Evidence regarding the diagnostic accuracy of the tuberculin skin test for patients with compromised immunity prior to initiating biologic treatment was available from one systematic review with meta-analyses (that included 16 relevant primary studies) and three additional diagnostic test accuracy studies. The results of these studies were inconsistent. The findings of the systematic review suggested that the coherence between the tuberculin skin test and interferon-gamma release assays was relatively high at 85%; however, the authors of the three diagnostic test accuracy studies concluded that the agreement between the two tests in their study populations was poor or fair.

One relevant non-randomized study was identified regarding the clinical utility of using the tuberculin skin test for patients with compromised immunity prior to initiating biologic treatment. This study was non-comparative and was descriptive in nature. The authors noted that the frequency of active tuberculosis infection was 6.8% during 55 months of follow-up after patients initiated biologic therapy, despite being screened for latent tuberculosis using the tuberculin skin test (those who tested positive for the tuberculin skin test were given prophylaxis with isoniazid).

Six guidelines that provided recommendations regarding the testing for latent tuberculosis infection in patients with compromised immunity prior to initiating biologic therapy were identified. All six of these guidelines recommended in favour of screening for latent tuberculosis in patients who are candidates for biologic therapy due to their increased risk for developing active tuberculosis. The tuberculin skin test, interferon-gamma release assays, chest x-ray, and chest computed tomography were proposed as viable diagnostic techniques; however, recommendations did not clearly indicate one was preferred over the others in all clinical scenarios. The underlying evidence informing these recommendations was generally of low quality or the recommendations were based solely on the expert opinion of the guideline development groups.

The limitations of the included literature (e.g., the lack of an established gold standard for detecting latent TB infection, the unclear generalizability to Canadian settings) should be considered when interpreting the findings of this report.

Methods

Literature Search Methods

This report is an update of a literature search strategy developed for a previous CADTH report.[9] For the current report, a limited literature search was conducted on key resources including MEDLINE, the Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases, Canadian and major international health technology agencies, as well as a focused internet search. No filters were applied to limit the retrieval by study type. The initial search was limited to English-language documents published between January 1, 2015 and June 16, 2020. For the current report, database searches were rerun on October 22, 2020 to capture any articles published since the initial search date. The search of major health technology agencies was also updated to include documents published since June 2020.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts identified from the literature search were reviewed and potentially relevant articles were retrieved *and* assessed for inclusion. Additionally, all references identified for inclusion in the main body or appendix of the previous CADTH report were retrieved for full-text assessment. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Patients with compromised immunity prior to starting biologic therapy (e.g., anti-TNF-α therapy, HIV treatments or other immune suppressants)
Intervention	Q1,2: TST Q3: TST or IGRA
Comparator	Q1: IGRA (using any reference standard) Q2: No comparator; IGRA; no screening test for latent TB infection Q3: Not applicable
Outcomes	Q1: Diagnostic accuracy (e.g., sensitivity, specificity) Q2: Clinical utility (e.g., diagnosis of TB, adverse events) Q3: Recommendations regarding the use of TST or IGRA for patients with compromised immunity prior to initiating biologic therapy (i.e., does this confirm a diagnosis of TB and lead to optimal treatment options)
Study Designs	Health technology assessments, systematic reviews, randomized controlled trials, non-randomized studies, and evidence-based guidelines

IGRA = interferon-gamma release assay; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to 2015. Systematic reviews in which all relevant studies were captured in other more recent or more comprehensive systematic reviews were excluded. Guidelines that had been updated since their original publication

were excluded if the newer version was included in this review. Primary studies retrieved by the search were excluded if they were captured in one or more included systematic reviews. Guidelines with unclear methodology were also excluded.

Critical Appraisal of Individual Studies

The included publications were critically appraised by one reviewer using the following tools as a guide: A MeaSurement Tool to Assess systematic Reviews 2 (AMSTAR 2)[10] for systematic reviews, the Downs and Black checklist[11] for non-randomized studies, the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2) checklist[12] for diagnostic test accuracy studies, and the Appraisal of Guidelines for Research and Evaluation (AGREE) II instrument[13] for guidelines. Summary scores were not calculated for the included studies; rather, the strengths and limitations of each included publication were described narratively.

Summary of Evidence

Quantity of Research Available

A total of 514 citations were identified from the electronic literature search conducted as part of the previous CADTH report.[9] Following screening of titles and abstracts, 503 of these citations were excluded. An additional 139 citations were identified from the updated electronic literature search conducted for the current report, of which 124 citations were excluded following screening of titles and abstracts. Eight potentially relevant publications were retrieved from the grey literature searches for full-text review. Of these 34 potentially relevant articles, 23 publications were excluded for various reasons, while 11 publications met the inclusion criteria and were included in this report. These comprised one systematic review with meta-analyses,[14] one non-randomized study,[15] three diagnostic test accuracy studies,[16-18] and six evidence-based guidelines.[19-24] Appendix 1 presents the PRISMA[25] flowchart of the study selection.

Additional references of potential interest are provided in Appendix 5.

Summary of Study Characteristics

One relevant systematic review with meta-analyses,[14] one non-randomized study,[15] three diagnostic test accuracy studies,[16-18] and six evidence-based guidelines[19-24] were identified for inclusion in this review. No relevant health technology assessments or randomized controlled trials were identified. Detailed study characteristics of included publications are provided in Appendix 2, Table 2, Table 3, and Table 4.

Study Design

The Alrajhi et al. (2020)[14] systematic review included primary studies published between June 2011 and April 2018. A total of 16 studies (one RCT and 15 observational cohort studies) were included in the review.

One non-randomized study[15] was identified regarding the clinical utility of using the TST for patients with compromised immunity prior to initiating biologic treatment. The study was a descriptive observational study that was conducted at a single centre. Data were collected between January 2013 and August 2017.

All three diagnostic test accuracy studies[16-18] enrolled consecutive patients and assessed individuals with two index tests (i.e., TST and IGRAs). One study[17] also examined patients with chest computerized tomography, which the authors considered a reference test for the purposes of calculating diagnostic parameters. One study[16] was prospective and was conducted at multiple centres and two studies[17,18] were retrospective, single-centre studies.

All six evidence-based guidelines[19-24] were informed by systematic reviews of the literature and included recommendations that were drafted using various consensus-generating methods. The Smith et al. (2020)[19] guidelines were developed by the British Association of Dermatologists. The quality of the evidence informing the recommendations was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) system where evidence was ranked as being of high, moderate, low, or very low quality. Recommendations were classified as strong or weak, based on the interpretation of the evidence by the developing group. The guidelines by Ooi et al. (2019),[20] which were developed on behalf of the Asia Pacific Association of Gastroenterology Working Group on Inflammatory Bowel Disease and Asian Organization for Crohn's and Colitis, graded recommendations between A (highest) and E (lowest) according to the Canadian Task Force on the Periodic Health Examination criteria. Consensus statements were drafted using a modified Delphi process. The guidelines by Rodríguez-Jiménez et al. (2018)[21] provided recommendations that were classified as strong, moderate, or weak, based on the quality of the evidence informing the recommendations, which was graded between Grade I (highest) and Grade III (lowest). The fourth guideline[22] was authored by the Guideline Development Group of the Clinical Practice Guideline on Systemic Lupus Erythematosus, which included experts from many collaborating societies and the Spanish Ministry of Health. Their recommendations were classified based on the Scottish Intercollegiate Guidelines Network (SIGN) criteria. A grade between A (highest) and ✓ (lowest) was assigned to each recommendation. Scientific evidence was classified between 1++ (highest quality) and 4 (lowest quality). Recommendations related to diagnostic tests were graded between A (highest grade) and D (lowest grade), with the underlying evidence being graded between Ia (highest quality) and IV (lowest quality). The Santin et al. (2016)[23] guidelines, which were developed by the Mycobacteria Study Group of the Spanish Society of Infectious Diseases and the Clinical Microbiology and the Spanish Society of Respiratory Diseases and Thoracic Surgery, classified recommendations as strong or weak based on evidence that was given a quality between high and very low using the GRADE criteria. The Singh et al. (2015)[24] guidelines were developed by the American College of Rheumatology. Their recommendations were classified as being strong or conditional based on the underlying evidence that was ranked as being of high, moderate, low, or very low quality (using the GRADE system).

Country of Origin

The included systematic review[14] was by authors in Canada. Primary studies included in the systematic review were conducted in Australia, Bulgaria, Canada, China, Denmark, Greece, Hungary, Italy, Portugal, Romania, Spain, Turkey, and the United States.

The non-randomized study[15] was conducted in Brazil and the three diagnostic test accuracy studies were conducted in Hong Kong,[18] Tunisia,[17] and Turkey.[16]

The guidelines were intended for use in Spain,[21-23] the United Kingdom,[19] the United States,[24] and for use throughout Asia.[20]

Patient Population

The systematic review[14] included primary studies that recruited adults (≥18 years of age) with inflammatory bowel disease prior to initiating anti-TNF medications. A total of 2,488 patients were included in the studies selected for inclusion in the review. Detailed patient characteristics, such as gender and disease severity, were not available in the systematic review.

The non-randomized study included 161 individuals who were diagnosed with rheumatoid arthritis, ankylosing spondylitis, juvenile idiopathic arthritis, or psoriatic arthritis and who received therapy with biologics. The mean age of participants was not reported and the proportion of participants who were female was 68.9%. The diagnostic test accuracy study by Girit et al. (2020)[16] recruited 57 children (< 18 years of age) who were diagnosed with immune-mediated inflammatory diseases and were scheduled to receive anti-TNF therapy. Children who had a history of TB treatment, latent TB prophylaxis, or prior anti-TNR drug use were excluded from the study. The mean age of participants was 12.4 years and the proportion of females was 59.6%. The study by Sellami et al. (2019)[17] included 105 adults (44 males, 61 females; mean age of 47.7 years) who were diagnosed with rheumatoid arthritis, spondyloarthritis, Still's disease, or chronic inflammatory bowel disease with an indication for biologic treatment. Patients who had a history of other autoimmune diseases, or who had already received an anti-TB chemoprophylaxis were excluded. The third diagnostic test accuracy study, conducted by So and colleagues,[18] included individuals who were diagnosed with rheumatoid arthritis, spondyloarthritis, or psoriatic arthritis and were considered candidates for biologic therapy. Those with active TB infection or a history of incomplete TB treatment were excluded. Participants' mean age was 49.3 years and the proportion of female participants was 68.4%.

The main populations covered by the guidelines were patients with psoriasis,[19] patients with ulcerative colitis or Crohn's disease,[20] patients with dermatologic conditions who are candidates for biologic therapy,[21] adults with systemic lupus erythematosus,[22] patients at risk for or suspected of having active TB,[23] and adults with rheumatoid arthritis. The intended users of all six guidelines were clinicians who care for these individuals.

Interventions and Comparators

The systematic review investigated the concordance between both types of new IGRA generations (QuantiFERON TB-2G or 3G) and the TST for the detection of latent TB.

The non-randomized study[15] screened for latent TB using TST (according to Mantoux method) and chest X-ray assessment. The three diagnostic test accuracy studies[16-18] assessed their study populations using both the TST and IGRAs. In the study by Girit et al. (2020),[16] screening was conducted using the TST and the T-SPOT.TB IGRA. Diagnostic parameters were calculated using the IGRA test as the index test and the TST at various cut-off values as the reference standard. Sellami et al. (2019)[17] screened for latent TB infection using the TST, IGRA (using the QuantiFERON-TB Gold Plus test), and chest computerized tomography. Both the TST and IGRA were considered index tests and chest computerized tomography was considered the reference standard tor their analyses. Participants of the study by So et al. (2017)[18] received screening with IGRA (using the ASACIR.TB) and the TST simultaneously.

All six evidence-based guidelines[19-24] made recommendations regarding the use of the TST or IGRAs to screen for latent TB prior to initiating biologic therapy.

Outcomes

The systematic review[14] evaluated test concordance between TST and IGRA for detecting latent TB.

The non-randomized study[15] assessed incidence of TB infection after treatment commencement with anti-TNF therapy and the number of patients who developed active TB after receiving prophylaxis (following a positive TST). All three diagnostic accuracy studies[16-18] reported on the agreement between the TST and IGRAs within their study populations. In addition, the study by Girit et al. (2020)[16] reported on the results of the screening tests before and after six months of anti-TNF therapy (i.e., TST and IGRA positivity rates), Sellami et al. (2019)[17] reported on various parameters of diagnostic accuracy (i.e., sensitivity, specificity, and accuracy), and So et al. (2017)[18] reported on the number of patients who developed active TB.

All six guidelines[19-24] drafted the recommendations relevant to the current report while considering the diagnostic accuracy of TST and IGRAs (e.g., specificity, sensitivity, predictive values) and the efficacy of chemoprophylaxis based on the diagnostics results.

Summary of Critical Appraisal

Additional details regarding the strengths and limitations of the included publications are provided in Appendix 3, Table 5, Table 6, Table 7, and Table 8.

Systematic Review

The included systematic review[14] was generally well-conducted according to AMSTAR 2 criteria. The review[14] had clearly defined research questions, objectives, and eligibility criteria. The study protocol was registered using an international database of prospectively registered systematic reviews (PROSPERO), decreasing the risk for selective reporting. Relevant literature was identified using a comprehensive literature search strategy that was performed in multiple databases (MEDLINE, Embase, and Cochrane Library databases). The electronic searches were supplemented with recursive searches, cross-referencing, and hand-searches. Key search terms and the dates of searches were provided, increasing the reproducibility of the literature search. The review[14] included a diagram that illustrated the article selection process and provided reasons for article exclusion. The methods for article selection, data extraction, and quality assessment were well-documented and were conducted in duplicate (disagreements were resolved through discussion and consensus, involving a third reviewer if required), decreasing the likelihood for inconsistency in these processes. The review authors described the included primary studies in adequate detail, summarizing their country or origin, patient population, study design, and main findings, and assessed their risk of bias using a satisfactory technique (i.e., the QUADAS-2). Meta-analyses were conducted using appropriate methods and statistical heterogeneity was assessed using the I^2 statistic. Finally, review authors disclosed their potential conflicts of interest (which were considered unlikely to have influenced the findings of the reviews) and reported on their sources of funding (no funding was received for this review).

As for the limitations of the review,[14] it was unclear which study designs were eligible for inclusion and the literature searching did not include a grey literature search, increasing the risk for missing relevant, non-indexed articles. Although reasons for excluding articles after full-text review were provided, the review did not include a list of these excluded studies. The review authors did not report on sources of funding for their included primary studies and the did not carry out an adequate investigation of publication bias.

Non-Randomized Study

The included non-randomized study was assessed as having some methodological strengths, mostly related to the reporting of results. The objectives, interventions, patient eligibility criteria, outcomes, and main findings were clearly described within the publication.[15] Patient characteristics, such as age, gender, clinical diagnosis, and duration of disease were provided. Due to the nature of the study (i.e., a retrospective descriptive observational study), no participants were lost to follow-up. Study participants, care providers, and health care settings appeared to be representative of the population and care setting of interest, increasing the external validity of the study. Both the sources of funding and potential conflicts of interest were disclosed and were unlikely to have influenced the findings of the study.

This non-randomized study[15] had significant methodological limitations. First and foremost, this descriptive observational study intended to evaluate the association between use of biologic agents and the frequency of TB infection a cohort of rheumatic patients. Because this study did not include a control group it did not provide comparative evidence on the clinical utility of the TST versus IGRA or no screening for latent TB prior to initiating biologic therapy, it simply provided a description of one center's experience in assessing rheumatic patients with the TST prior to initiating biologic therapy. Additionally, adverse events that may have been associated with the use of the TST were not reported (e.g., adverse events associated with isoniazid therapy for latent TB when a false positive result was received for the TST). There was no blinding of study participants or outcome assessors and this was a single-centre study conducted in the Brazil; the generalizability to the Canadian setting was unclear.

Diagnostic Test Accuracy Studies

There were several strengths common to the three diagnostic test accuracy studies,[16-18] including: 1) clearly described objectives, interventions, controls, inclusion and exclusion criteria, outcomes, participant characteristics (e.g., age, sex, clinical diagnosis, time since diagnosis, current and past medical therapy), and main findings, 2) participants appeared to have been recruited consecutively (rather than at random), 3) no studies used a case-control study design, 4) inappropriate participant exclusion criteria were avoided, 5) the cut-off values for a positive TST were pre-specified, 6) there was an appropriate time interval between the two diagnostic tests (i.e., they were given simultaneously), 7) study participants, care providers, and setting appear to be representative of the population and care setting of interest, and 8) all study authors declared that they had no potential conflicts of interest. The studies by Girit et al. (2020)[16] and Sellami et al (2019)[17] also reported on their sources of funding (no funding was received) and the studies by Sellami et al. (2019)[17] and So et al. (2017)[18] stated that the two diagnostic tests were interpreted by clinicians who were blinded to the result of the other test (i.e., the TST was interpreted without knowledge of the results for the IGRA).

As for methodological limitations, three studies[16-18] stated that there was no gold standard method for the diagnosis of latent TB infection, and therefore the authors used Cohen's kappa to evaluate the coherence and inter-rater agreement between the TST and the IGRA. While this method can evaluate the agreement between the two screening methods, it does not assess whether either test correctly classified the target condition. In addition, the studies by Girit et al. (2020)[16] and Sellami et al. (2019)[17] calculated diagnostic parameters (e.g., sensitivity, specificity, positive predictive value, negative predictive value) using the TST at various cut-off values and chest computerized tomography as reference

standards, respectively; however, these findings should be interpreted knowing that neither of these reference standards was considered a gold standard (and are therefore likely to inaccurately classify the target condition). It was unclear if the two diagnostic tests were interpreted without knowledge of the result of the other test in the study by So et al. (2017)[18] and Girit et al. (2020)[16] did not report on their sources of funding. The three studies[16-18] were conducted in Turkey, Tunisia, and Hong Kong, therefore the generalizability of their findings to the Canadian setting is unclear.

Evidence-Based Guidelines

The six evidence-based guidelines were of varying methodological quality,[19-24] based on assessments made using the AGREE II checklist.

All six guidelines[19-24] provided clear descriptions of their scope and purpose, including objectives, the range of clinical questions covered in the guidelines, the intended users, and target population. The final recommendations were easily identifiable, considered the health benefits, side effects, and risks, had an explicit link to the supporting evidence, and were written using language that was clear and unambiguous. In all cases, the guideline development groups appeared to include individuals from all relevant professional groups. For example, the Santin et al. (2016)[23] guidelines on the use of IGRA in the diagnosis of TB were drafted by a panel that included experts in infectious diseases, respiratory diseases, microbiology, pediatric preventive medicine, and a methodological expert. Three guidelines[19,22,24] specifically sought and incorporated the views and preferences of patients and the public as part of their development; however, three guidelines[20,21,23] did not. Systematic literature searches were used to identify evidence for consideration when developing recommendation in all six guidelines. Search strategies (e.g., search terms and search filters), databases searched, and the timing of the literature searches were clearly described in four guidelines.[19,22-24] The authors of two guidelines[20] appeared to use a systematic approach to identifying evidence but they did not report on these methods in sufficient detail. The guidelines by Ooi et al. (2019)[20] included information on the databases searched but did not provide key search terms, dates the searches were performed, or the eligible study designs. Similarly, Rodríguez-Jiménez et al. (2018)[21] stated they searched for publications in Medline and Cochrane databases up until July 2017, but provided no additional information on their literature search. Two guidelines[19,24] provided a detailed description of their methods used for selecting articles and stated that the article selection process was done in duplicate. Two guidelines[22,23] described their methods for selecting articles in adequate detail (i.e., they provided inclusion and exclusion criteria) but they did not state if the selection process was done in duplicate. Two guidelines[20,21] did not provide any details relating to their methods of evidence selection. In all cases, the guidelines[19-24] included some description of how the recommendations were formulated, which were made using various consensus-generating procedures (e.g., modified Delphi process). Two guidelines[21,23] stated that recommendations were established by consensus but provided no additional description of the processes used. The descriptions of the methods for recommendation formulation were considered sufficient in the four remaining guidelines.[19,20,22,24] Four guidelines[19,22-24] were externally reviewed by experts prior to their publication and included a procedure for updating the guideline in the future (the other two guidelines[20,21] did not have either feature). As for applicability to clinical practice, all six guidelines[19-24] included advice or various tools for clinicians (e.g., clinical decision algorithms, patient handouts) and considered the resource implications of implementing the recommendations. Two guidelines[19,24] also included criteria for monitoring or auditing the implementation of the guidelines into practice. Editorial independence was established for

five guidelines,[19,20,22-24] where competing interests were disclosed by the guideline development group and the views of the funding body did not appear to have any influence on the content of the guidelines. Rodríguez-Jiménez et al. (2018)[21] included a description of potential conflicts of interest; however, their sources of funding were not disclosed, making it unclear if the funders' views had any impact on the content of the guidelines. Finally, it should be noted that these guidelines were developed for use in the United Kingdom,[19] Spain,[21-23] the United States,[24] and various Asia-Pacific countries;[20] therefore, the generalizability of the recommendations to the Canadian context is unclear.

Summary of Findings

The overall findings of the included studies are highlighted below. Detailed summaries of the main findings are available in Appendix 4, Table 9, Table 10, and Table 11.

Diagnostic Accuracy of the Tuberculin Skin Test Prior to Initiating Biologic Treatment

Diagnostic parameters

Evidence regarding the diagnostic accuracy of the TST versus IGRAs for the detection of latent TB was available from one systematic review with meta-analyses[14] and two diagnostic test accuracy studies.[16,18]

The systematic review[14] included a meta-analysis that pooled data from 16 relevant primary studies. The authors reported a pooled concordance between the TST and IGRAs of 85% (95% confidence interval [CI] = 81% to 88%; P value =0.01; I^2 = 76%). Additional meta-analyses were conducted to examine the odds ratio of testing positive by IGRA or by the TST in immunosuppressed populations versus non-immunosuppressed populations. There were no statistically significant differences in the odds ratio of testing positive by the TST between immunosuppressed populations and non-immunosuppressed populations when using a random effects model (OR = 1.14; 95% CI = 0.61 to 2.12; P value = 0.69; I^2 = 42%). Similarly, there were no statistically significant differences in the odds ratio of testing positive by IGRAs in immunosuppressed populations versus non-immunosuppressed populations when a random effects model was applied (OR = 0.57; 95% CI = 0.31 to 1.03; P value = 0.06; I^2 = 32%); however, when a fixed effects model was applied, which the authors considered appropriate given the low heterogeneity, the odds ratio of testing positive with IGRAs in immunosuppressed populations versus immunosuppressed populations reached statistical significance (OR = 0.57; 95% CI = 0.37 to 0.89; P value = 0.01; I^2 = 32%).

The diagnostic test accuracy study by Girit and colleagues[16] reported on the number of patients who tested positive for latent TB with the TST at various cut-off points (i.e., ≥5 mm, 5 to 9 mm, 10 to 14 mm, and ≥15 mm) and IGRAs prior to initiating biologic therapy and after six months of treatment with biologic therapy (N = 57). Prior to initiating biologic therapy, the agreement between IGRA and the TST (using the standard ≥5 mm cut-off) was poor (Cohen's kappa [κ] = 0.176). The sensitivity was 60.00%, the specificity was 75.00%, the positive predictive value was 18.75%, the negative predictive value was 95.12%, and the accuracy was 73.68%. Agreement between the two tests (assessed using Cohen's kappa) using other TST cut-offs and after six months of biologic therapy ranged between 0 and 0.339, indicating no agreement to fair agreement (in most cases agreement was considered poor). Additional parameters, such as sensitivity, specificity, positive predictive value, negative predictive value, were reported and the findings are described in Appendix 4, Table 10. The authors concluded that there was poor coherence between TST and IGRA

in patients who are scheduled for anti-TNF therapies. Sellami et al. (2019)[17] determined the agreement between IGRA and TST was low (Cohen's kappa [κ] = 0.08; 95% CI = 0.003 to 0.162) when using a 10 mm TST cut-off in 95 patients with chronic immune-mediated inflammatory diseases prior to biologic therapy. The diagnostic parameters of the test tests were also calculated. When using chest computerized tomography as the reference standard, sensitivity, specificity, and accuracy for IGRA were 26.14%, 100%, and 75.3%, respectively. These values compared to a sensitivity of 73.9%, a specificity of 80.6%, and an accuracy of 73.6% for TST using a 10 mm cut-off. Similar results were reported by So et al. (2017).[18] Within their study population (N = 38), agreement between IGRA and TST was moderate when using a 5 mm cut-off (Cohen's kappa [κ] = 0.47; standard error [SE] = 0.13) and was fair when using a 10 mm cut-off (Cohen's kappa [κ] = 0.39; SE = 0.16). The authors also noted that agreement between the two tests using a 10 mm cut-off in patients on prednisolone (N = 19) had slight agreement (Cohen's kappa [κ] = 0.066; SE = 0.23), and moderate agreement in those not on prednisolone (Cohen's kappa [κ] = 0.57; SE = 0.16). When using a 5 mm cut-off, the two testing strategies had substantial agreement (Cohen's kappa [κ] = 0.69; SE = 0.16) in patients not on prednisolone (N = 19).

Clinical Utility of the Tuberculin Skin Test Prior to Initiating Biologic Treatment

Diagnosis of TB

Information regarding the clinical utility of the TST prior to initiating biologic treatment was available from one non-randomized, descriptive observational study.

The study included 161 rheumatic patients who were assessed for latent TB using the TST prior to starting biologic therapy. The TST was positive in 19.25% (N = 31) of these patients. Those who tested positive were given latent TB prophylaxis with isoniazid (300 mg daily) for six months. Biologic treatment was started only after the patient received isoniazid for at least two months. Throughout the course of the study, 11 of the 161 patients (6.8%) developed active TB (four of these individuals tested positive for the TST and developed active TB despite prophylactic treatment). While these findings were descriptive, rather than comparative (i.e., there was no control group), the authors concluded that the frequency of TB infection during 55 months of follow-up was 6.8% among 161 consecutive rheumatic patients receiving biologics.

Guidelines

Recommendations regarding the screening of latent TB infection prior to biologic therapy

Six evidence-based guidelines[19-24] were identified that made recommendations regarding the testing for latent TB infection in patients with compromised immunity prior to initiating biologic therapy.

The guidelines by Smith et al. (2020)[19] gave a weak recommendation in favour of screening for latent TB with an IGRA alone, or with an IGRA and concurrent Mantoux test while considering individual's risk factors for TB when interpreting results in patients who are going to receive biologic therapy for psoriasis. This recommendation was based on consensus following consideration of guidelines published by the National Institute for Health and Care Excellence (NICE).

Ooi et al. (2019)[20] recommend routine screening for latent TB infection in those with ulcerative colitis and Crohn's disease according to local practice before initiating biologic treatment which may include chest X-ray, chest computer tomography, IGRA and/or TST.

This recommendation was classified as having good evidence to support the statement, which was from well-designed cohort or case-control studies.

Guidelines by Rodríguez-Jiménez et al. (2018)[21] included a total of ten relevant recommendations regarding the screening for TB infection in dermatology patients who are candidates for biologic therapy. Eight strong recommendations, based on expert opinion in descriptive studies or in clinical experience, were provided: 1) all candidates for biological therapies should be examined for latent TB infection as these patients are at high risk for developing TB, 2) the risk of TB appears to depend on the anti-TNF agent used (infliximab and adalimumab are those at the highest risk), 3) diagnostic methods for latent TB infection include examination of patients medical records and contacts with those who have active TB, chest X-ray and computer tomography for detecting old TB lesions, and simultaneous performance of IGRA and TST (with a positive result in either of these indicating latent TB infection), 4) patients with immune-mediated inflammatory diseases have a higher rate of false negatives with TST and IGRA, 5) IGRA should be performed before TST, due to the booster effect detected for IGRAs, 6) indeterminate results should always be confirmed with a second measurement, 7) negativity in the TST and IGRA does not rule out latent TB infection (i.e., these diagnostic tests have a considerable false negative rate), and 8) preventive treatment is recommended in all candidates for therapy who present with suspicion of latent TB infection, after excluding active TB, in any of the diagnostic tests. A moderate recommendation, based on expert opinion in descriptive studies or in clinical experience, states that the specificity and sensitivity of the two IGRA techniques (i.e., QuantiFERON-TB-Gold in tube and T-SPOT.TB) are similar in patients with immune-mediated inflammatory diseases; however, the sensitivity of T.SPOT.TB is somewhat higher in patients treated with corticosteroids and should be considered for use in these patients. Finally, the guideline included a weak recommendation suggesting that there is no evidence that repeating the TST increases the sensitivity of the test in patients with immune-mediated inflammatory diseases but that it does reduce specificity; therefore, with the use of IGRA, the TST is not recommended.

The guidelines by the GDG of the CPG on SLE (2016)[22] included four relevant recommendations regarding the diagnosis of latent TB infection in patients with systemic lupus erythematosus. Three recommendations were based on clinical experience and the consensus of the drafting team, 1) screening for TB should be adapted to the clinical situation and the individual risk factors of each patient, 2) patients who are going to be submitted to immunosuppressive treatment should be screened for TB, especially when the treatment involves high doses of glucocorticoids or biological therapies, regardless of the existence of risk factors, and 3) the TST is the test of choice to detect TB due to its sensitivity using the standard cut-off point (5 mm); however, previous BCG vaccination and/or immunosuppression, could make IGRAs (specifically, the QuantiFERON Gold Test) more reliable test for detecting latent infection. A grade D recommendation suggests that when a negative result is received for the TST, a second test should be conducted one week later to induce the immunological memory (booster effect) since false negatives are more frequent in the elderly and in immunosuppressed patients.

Santin et al. (2016)[23] gave a weak recommendations for using both the TST and an IGRA to screen for TB infection in patients with chronic inflammatory diseases before starting biological therapy. This recommendation was informed by low to very low-quality evidence from eight observational studies.

The guidelines by Singh et al. (2015)[24] endorsed a screening algorithm for patients who are candidates for biologics or tofacitinib. The algorithm indicated that initial screening should be performed with the TST or IGRA, with a preference for IGRA in patients with a history of BCG vaccination. Rescreening was recommended if TB risk factors are present. The strength of this recommendation and the quality of the underlying evidence was unclear.

Limitations

While the included systematic review[14] was generally well-conducted according to AMSTAR II criteria, the underlying evidence from the included primary studies was rated as being of moderate to high risk of bias in many cases (as assessed by the authors of the systematic review). Common methodological limitations from the primary studies included concerns related to the selection of the index test and reference standard, the flow and timing of testing, and the methods used for patient selection. Any quality issues from the primary studies cause uncertainty in the findings presented in the systematic review.

Studies assessing the diagnostic accuracy of the TST and IGRAs were unable to compare the results of the screening tests to an established gold standard technique for diagnosing latent TB. High agreement between these two tests does not necessarily indicate that they are both effective in detecting latent TB.

No studies assessing the comparative clinical utility of the TST and IGRAs for detecting latent TB were identified. It is unclear which of these two tests may result in improved clinical outcomes in immunosuppressed patients prior to biologic therapy.

With the exception of a single primary study included in the systematic review,[2] none of the studies[14-18] summarized in this report were conducted in Canada. Similarly, none of the included guidelines[19-24] were intended for users in Canada. Therefore, the generalizability of the findings from the included literature and the applicability of the recommendations from the included guidelines to Canadian settings are unclear.

Conclusions and Implications for Decision or Policy Making

This review was comprised of one systematic review with meta-analyses,[14] one non-randomized study,[15] and three diagnostic test accuracy studies.[16-18] Of these studies, the systematic review and the three diagnostic test accuracy studies[16-18] addressed the comparative diagnostic accuracy of the TST versus IGRAs for detecting latent TB in patients with compromised immunity prior to initiating biologic treatment while the non-randomized study[15] assessed the clinical utility of providing the TST for detecting latent TB in patients with compromised immunity prior to initiating biologic treatment. In addition, six evidence-based guidelines[19-24] that provided recommendations regarding the testing for latent TB infection in patients with compromised immunity prior to initiating biologic therapy were included.

Based on the evidence summarized in this report, the agreement between the TST and IGRAs appears to inconsistent. The systematic review[14] suggested that concordance between the two tests was relatively high (85%), but the three diagnostic test accuracy studies[16-18] found agreement between the two tests to be poor. The inconsistency in studies assessing the performance of these diagnostic tests was echoed in the evidence-based guidelines. All six guidelines recommended in favour of screening for latent TB infection prior to starting biologic therapy; however, there was rarely any clear guidance suggesting one screening test should be favoured over another. One guideline[21] even suggested that

negativity in the both the TST and IGRA does not rule out latent TB infection. In most cases, screening with both tests and conducting repeat screening when appropriate was the recommended approach, although there were certain clinical scenarios where one screening test was preferred over another (e.g., IGRAs may be preferred in patients with a history of Bacillus Calmette–Guérin vaccination).

Although this review was intended as an upgrade to a previously published report,[9] one systematic review[6] and three evidence-based guidelines[26-28] included in the predecessor report were excluded in this review following full-text assessment. The systematic review[6] and two guidelines[26,28] were excluded as they did not include recommendations specific to testing for latent TB prior to initiating biologic therapy and one guideline[27] was excluded as an updated version of the guideline was identified and included in the current report.

The limitations of the included literature (e.g., the lack of an established gold standard for detecting latent TB infection, the unclear generalizability to Canadian settings) should be considered when interpreting the results. Further research investigating the diagnostic accuracy and clinical utility of the TST and IGRAs, especially once a gold standard technique has been established for the detection of latent TB, would help reduce the uncertainty regarding the used of these diagnostic tests in patients with compromised immunity prior to initiating biologic therapy.

References

1. World Health Organization. Tuberculosis. 2020; https://www.who.int/news-room/fact-sheets/detail/tuberculosis, 2020 Nov 24.
2. Government of Canada. Tuberculosis: Monitoring. 2019; https://www.canada.ca/en/public-health/services/diseases/tuberculosis/surveillance.html, 2020 Nov 24.
3. World Health Organization. Tuberculosis profile: Philippines. 2020; https://worldhealthorg.shinyapps.io/tb_profiles/?_inputs_&lan=%22EN%22&iso2=%22PH%22, 2020 Nov 24.
4. Centers for Disease Control and Prevention. Tuberculosis (TB). 2014; https://www.cdc.gov/tb/publications/factsheets/general/ltbiandactivetb.htm, 2020 Nov 24.
5. Rabahi MF, Silva Júnior JLRd, Ferreira ACG, Tannus-Silva DGS, Conde MB. Tuberculosis treatment. *Jornal brasileiro de pneumologia : publicacao oficial da Sociedade Brasileira de Pneumologia e Tisilogia.* 2017;43(6):472-486.
6. Hasan T, et al. Screening and prevention for latent tuberculosis in immunosuppressed patients at risk for tuberculosis: a systematic review of clinical practice guidelines. *BMJ Open.* 2018;8(9):e022445. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6144320/ Accessed 2020 Oct 7.
7. Mancuso JD, Bernardo J, Mazurek GH. The Elusive "Gold" Standard for Detecting Mycobacterium tuberculosis Infection. *American Journal of Respiratory and Critical Care Medicine.* 2013;187(2):122-124.
8. Mwaba P, Chakaya JM, Petersen E, Wejse C, Zumla A, Kapata N. Advancing new diagnostic tests for latent tuberculosis infection due to multidrug-resistant strains of Mycobacterium tuberculosis — End of the road? *Int J Infect Dis.* 2020;92:S69-S71.
9. Kumar D, Argáez C. Latent Tuberculosis Infection Testing in People with Compromised Immunity Prior to Biologic Therapy: Diagnostic Accuracy, Clinical Utility and Guidelines. *(CADTH rapid response report: summary of abstracts)*. Ottawa: CADTH; 2020: https://cadth.ca/sites/default/files/pdf/htis/2020/RB1502%20TST%20for%20Immunocompromised%20Final.pdf. Accessed 2020 Nov 24.
10. Shea BJ, Reeves BC, Wells G, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. *Bmj.* 2017;358:j4008. http://www.bmj.com/content/bmj/358/bmj.j4008.full.pdf. Accessed 2020 Jan 7.
11. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. *J Epidemiol Community Health.* 1998;52(6):377-384. http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf. Accessed 2020 Jan 7.
12. Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med.* 2011;155(8):529-536.
13. Agree Next Steps Consortium. The AGREE II Instrument. Hamilton (ON): AGREE Enterprise; 2017: https://www.agreetrust.org/wp-content/uploads/2017/12/AGREE-II-Users-Manual-and-23-item-Instrument-2009-Update-2017.pdf. Accessed 2020 Jan 7.
14. Alrajhi S, Germain P, Martel M, et al. Concordance between tuberculin skin test and interferon-gamma release assay for latent tuberculosis screening in inflammatory bowel disease. *Intestinal Research.* 2020;18(3):306-314.
15. Chaer FGG, de Lucena Valim JM, Reis RC, Klautau GB, de Souza BDB. Use of biologic agents and risk of tuberculosis in Brazil, a tuberculosis high-burden country. *Drugs Context.* 2020;9:212598.
16. Girit S, Ayzit Atabek A, Şenol E, et al. Screening for Latent Tuberculosis in Children With Immune-mediated Inflammatory Diseases Treated With Anti-tumor Necrosis Factor Therapy: Comparison of Tuberculin Skin and T-SPOT Tuberculosis Tests. *Arch.* 2020;35(1):20-28.
17. Sellami M, et al. Screening for latent tuberculosis infection prior to biologic therapy in patients with chronic immune-mediated inflammatory diseases (IMID): Interferon-gamma release assay (IGRA) versus tuberculin skin test (TST). *Egypt Rheumatol.* 2019 Jul;41(3):225-250. https://www.sciencedirect.com/science/article/pii/S1110116418301352 Accessed 2020 Oct 7.
18. So H, Yuen CS, Yip RM. Comparison of a commercial interferon-gamma release assay and tuberculin skin test for the detection of latent tuberculosis infection in Hong Kong arthritis patients who are candidates for biologic agents. *Hong Kong medical journal = Xianggang yi xue za zhi.* 2017;23(3):246-250.
19. Smith CH, Yiu ZZN, Bale T, et al. British Association of Dermatologists guidelines for biologic therapy for psoriasis 2020: a rapid update. *British Journal of Dermatology.* 2020;183(4):628-637.
20. Ooi CJ, et al. Best practices on immunomodulators and biologic agents for ulcerative colitis and Crohn's disease in Asia. *Intestinal Res.* 2019;17(3):285-310. https://www.irjournal.org/journal/view.php?doi=10.5217/ir.2019.00026 Accessed 2020 Oct 7.
21. Rodríguez-Jiménez P, Mir-Viladrich I, Chicharro P, et al. Prevention and treatment of tuberculosis infection in candidates for biologic therapy: A multidisciplinary consensus statement adapted to the dermatology patient. *Actas Dermosifiliogr.* 2018;109(7):584-601.
22. Guideline Development Group of the Clinical Practice Guideline on Systemic Lupus Erthematosus. Clinical practice guideline on systemic lupus erythematosus. 2016: https://portal.guiasalud.es/wp-content/uploads/2018/12/GPC_549_Lupus_SESCS_compl_en.pdf
23. Santin M, García-García JM, Domínguez J. Guidelines for the use of interferon-γ release assays in the diagnosis of tuberculosis infection. *Enfermedades infecciosas y microbiologia clinica.* 2016;34(5):303.e301-313.
24. Singh JA, et al. 2015 American College of Rheumatology Guideline for the treatment of rheumatoid arthritis. *Arthr Care Res.* 2015. https://www.rheumatology.org/Portals/0/Files/ACR%202015%20RA%20Guideline.pdf Accessed 2020 Oct 7.

25. Liberati A, Altman DG, Tetzlaff J, et al. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *Bmj.* 2009;339:b2700.

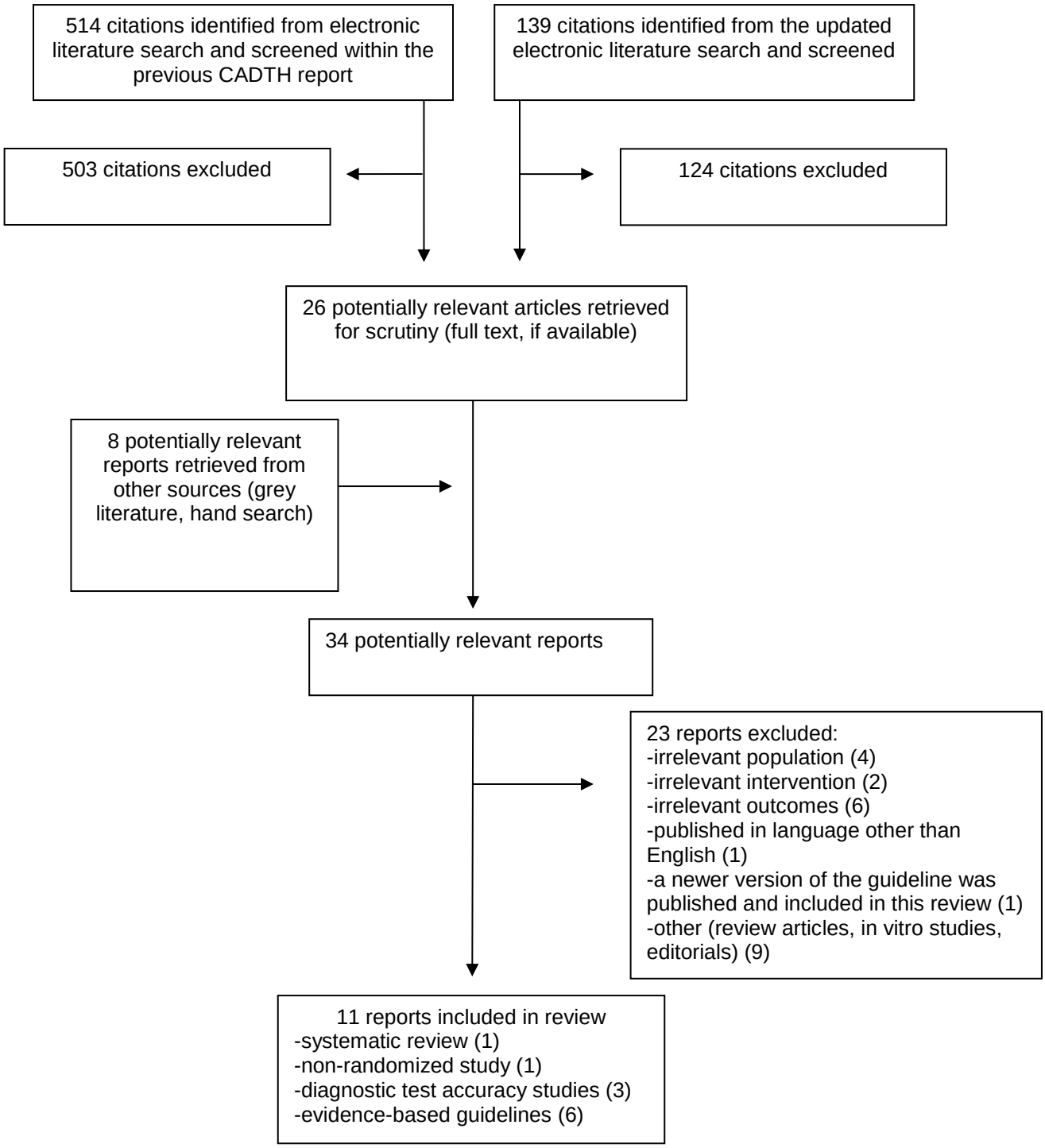
26. Singapore Ministry of Health. Prevention, Diagnosis, and Management of Tuberculosis. *(HOH clinical practice guidelines)* 2016; https://www.moh.gov.sg/docs/librariesprovider4/guidelines/moh-tb-cpg-full-version-for-website.pdf Accessed 2020 Oct 7.

27. British Association of Dermatologists. Guidelines for biologic therapy for psoriasis. 2017 Apr; https://www.bad.org.uk/shared/get-file.ashx?id=5835&itemtype=document Accessed 2020 Oct 7.

28. National Institute for Health and Care Excellence. Tuberculosis. *(NICE guideline NG33)* 2016; https://www.nice.org.uk/guidance/ng33 2020 Oct 7.

29. Shahidi N, Fu YT, Qian H, Bressler B. Performance of interferon-gamma release assays in patients with inflammatory bowel disease: a systematic review and meta-analysis. *Inflamm Bowel Dis.* 2012;18(11):2034-2042.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2: Characteristics of the Included Systematic Review

Study citation, country, funding source	Objective, study designs, search strategy, numbers of primary studies included, quality assessment tool	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
Alrajhi et al. (2020)[14] Canada **Funding source**: No financial support was received for this research	**Objective**: To assess test concordance between TSTs and IGRAs for latent TB screening in patients with inflammatory bowel disease prior to initiating anti-TNF medications. **Study designs**: Not explicitly stated, but the final analysis included prospective cohorts, retrospective cohorts, and RCTs. **Search strategy**: Electronic searches were conducted in MEDLINE, Embase, and Cochrane Library databases for articles published between June 2011 to April 2018. The June 2011 date was selected as it corresponded to a previously published systematic review and meta-analysis[29] on the same topic. The electronic searches were supplemented with recursive searches, cross-referencing, and subsequent hand-searches. **Number of primary studies included**: A total of 16 studies (1 prospective RCT, 13 prospective observational cohort studies, 2 retrospective observational cohort studies) were included in the review. **Quality assessment tool**: The quality of primary studies was assessed using QUADAS-2.	Adults with inflammatory bowel disease who were candidates for treatment with anti-TNF therapies.	The systematic review included studies that assessed the concordance between IGRAs and TSTs for the detection of latent TB.	**Primary outcome**: - Concordance between TSTs and IGRAs in inflammatory bowel disease patients being screened for latent TB infection prior to anti-TNF exposure **Secondary outcome**: - Effects of immunosuppressive therapy (defined as exposure to oral corticosteroids of more than 5 mg daily for at least 4 weeks, thiopurines, methotrexate or cyclosporine) on the results of TSTs and IGRAs prior to anti-TNF exposure **Follow-up**: NR.

IGRA = interferon-gamma release assay; NR = not reported; QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies-2; RCT = randomized controlled trial; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Table 3: Characteristics of Included Primary Clinical Studies

Study citation, country, funding source	Objective and study design	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
Non-randomized study				
Chaer et al. (2020)[15] Brazil **Funding source**: Funding for editorial assistance received from UCB Brazil.	**Objective**: To evaluate the frequency of TB infection in patients with rheumatoid arthritis, ankylosing spondylitis, juvenile idiopathic arthritis, and psoriatic arthritis after the use of biologic agents. **Study design**: Single centre observational study.	**Inclusion criteria**: Individuals who were diagnosed with rheumatoid arthritis, ankylosing spondylitis, juvenile idiopathic arthritis, or psoriatic arthritis according to ACR criteria and who received biological therapy. **Excluded**: No specific exclusion criteria were listed. **Number of participants**: 161 (113 with rheumatoid arthritis, 35 with ankylosing spondylitis, 5 with juvenile idiopathic arthritis, and 10 with psoriatic arthritis). **Mean age**: NR. **Gender**: 68.9% female.	All participants were screened for latent TB using TST (according to Mantoux method), chest X-ray assessment, and history of exposure to TB prior to anti-TNF treatments.	**Clinical outcomes**: - Incidence of TB infection after treatment commencement with anti-TNF therapy - Number of patients who developed active TB after receiving prophylaxis (following positive TST) **Follow-up**: - 55 months
Diagnostic test accuracy studies				
Girit et al. (2020)[16] Turkey **Funding source**: The authors received no financial support for this work.	**Objective**: To investigate the coherence between IGRA (T-SPOT.TB) and TST with different cut-off values in screening latent TB infection prior to and after six months of treatment with anti-TNF therapy. **Study design**: Multi-centre, prospective diagnostic test accuracy study.	**Inclusion criteria**: Children (<18 years of age) who were diagnosed with immune-mediated inflammatory diseases and were scheduled to receive anti-TNF therapy. **Excluded**: Those with a history of TB treatment, latent TB prophylaxis, or prior anti-TNF drug use. **Number of participants**: 57 (45 had juvenile idiopathic arthritis, 5 had systemic sclerosis, 4 had spondyloarthropathies, 3 had SLE) **Mean age**: 12.4 (SD = 3.9) years. **Gender**: 59.6% female	All participants received latent TB screening with IGRA and TST prior to initiation of anti-TNF therapy. Participants were then re-screened with both methods after 6 months of anti-TNF therapy.	**Clinical outcomes**: - Results of the screening tests before and after 6 months of anti-TNF therapy (i.e., TST and IGRA positivity rates) - Coherence between the two screening methods **Follow-up**: - 6 months

Study citation, country, funding source	Objective and study design	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
Sellami et al. (2019)[17] Tunisia **Funding source**: The authors received no financial support for this work.	**Objective**: To evaluate the agreement between IGRA and the TST in detecting latent TB infection in patients with chronic immune-mediated inflammatory diseases prior to biologic therapy. **Study design**: Single-centre, retrospective diagnostic test accuracy study.	**Inclusion criteria**: Adults (≥18 years of age) who were diagnosed with rheumatoid arthritis (using the 2010 ACR or European League Against Rheumatism criteria), spondyloarthritis (using the Assessment of SpondyloArthritis international Society classification criteria), or chronic inflammatory bowel disease (i.e., Crohn's disease or ulcerative colitis) with an indication for biologic treatment. Two patients with Still's disease refractory to treatment with corticosteroids and methotrexate were also included. **Excluded**: Those who were under the age of 18, had a history of other autoimmune diseases, or who had already received an anti-TB chemoprophylaxis. **Number of participants**: 105. **Mean age**: 47.7 (SD = 14.4) years. **Gender**: 58.1% female.	All participants were screened for active TB using medical history, clinical examination, chest radiography, and sputum examination. Latent TB infection was screened for using TST or/and IGRA. IGRA testing was performed using the QuantiFERON-TB Gold Plus.	**Clinical outcomes**: - Agreement between TST and IGRA - Diagnostic accuracy, using chest CT as the reference standard (e.g., sensitivity, specificity, accuracy) **Follow-up**: - Mean follow-up was 4.8 (SD = 1.4) years
So et al. (2017)[18] Hong Kong **Funding source**: NR.	**Objective**: To evaluate the concordance between IGRA and TST in the diagnosis of latent TB in patients with arthritic disease prior to treatment with biologic therapies. **Study design**: Single-centre, retrospective diagnostic test accuracy study.	**Inclusion criteria**: Individuals who were diagnosed with rheumatoid arthritis (using ACR or European League Against Rheumatism criteria), spondyloarthritis (using the Assessment of SpondyloArthritis international Society classification criteria), or psoriatic arthritis (using the Classification Criteria for Psoriatic Arthritis). Participants were included if they were considered candidates for biologic agents. **Excluded**: Those were had active TB infection, a history of incomplete TB treatment, or no measured induration. **Number of participants**: 38 (24 had rheumatoid arthritis 24, 12 had	All participants received screening with IGRA (ASACIR.TB) and TST simultaneously, prior to treatment with biologic therapy.	**Clinical outcomes**: - Results of the screening tests - Agreement between TST and IGRA - Number of patients who developed active TB **Follow-up**: - 1 year

Study citation, country, funding source	Objective and study design	Population characteristics	Intervention and comparator(s)	Clinical outcomes, length of follow-up
		spondyloarthritis, and 2 had psoriatic arthritis). **Mean age**: 49.3 (SD = 13.7) years. **Gender**: 68.4% female.		

ACR = American College of Rheumatology; CT = computerized tomography; IGRA = interferon-gamma release assay; NR = not reported; SD = standard deviation; SLE = systemic lupus erythematosus; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Table 4: Characteristics of Included Guidelines

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
Smith et al. (2020)[19] United Kingdom **Funding source**: The work was supported by the National Institute for Health Research (NIHR), a Medical Research Council (MRC) Clinical Academic Research Partnership award, a NIHR Clinician Scientist award, and the NIHR Oxford Biomedical Research Centre.	**Scope**: To provide up-to-date, evidence-based recommendations on the use of biologic therapies (e.g., adalimumab, etanercept, certolizumab pegol, infliximab, ustekinumab, ixekizumab, secukinumab, brodalumab, guselkumab, risankizumab, tildrakizumab) in individuals (of all ages) with psoriasis. This guideline updated a previously published guideline[27] that was also identified in our literature searches. **Developing institutions**: British Association of Dermatologists. **Intended users**: Clinicians who care for individuals with psoriasis. **Target population**: Adults, children, and young people with psoriasis.	**Evidence collection**: Systematic literature searches were conducted using PubMed, MEDLINE, Embase, and Cochrane databases to identify key articles relevant to each research question. Relevant study designs included SRs, RCTs, and observational studies. Search strategies varied by database, but there were generally no limits by publication date. Literature searches were conducted on September 7th, 2018. **Evidence selection**: Literature retrieved in the electronic searches was screened against the inclusion and exclusion criteria by two reviewers. **Evidence synthesis**: Meta-analyses and network meta-	The guideline development group formulated recommendations based on their interpretation of the available evidence (including results of the meta-analyses and quality assessments), taking into account the balance of benefits, harms and costs between different courses of action. When clinical evidence for a particular research question was poor or lacking, recommendations were based on expert opinion. A group consensus process was used to finalize recommendations.	**Recommendation grading system**: Recommendations were classified as being "strong" or "weak", based on the underlying evidence that was ranked as being of high, moderate, low, or very low quality (using the GRADE system). - Strong recommendation (↑↑): benefits of the intervention outweigh the risks; most patients would choose the intervention while only a small proportion would not; for clinicians, most of their patients would receive the intervention; for policymakers, it would be a useful performance indicator - Weak recommendation (↑): risks and benefits of the intervention are finely balanced; many patients would choose the intervention but many would not; clinicians would need to consider the pros and cons for the patient in the context of the evidence; for policymakers, it would be a poor performance

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
		analyses were conducted using Review Manager (RevMan) 5.3 to combine the data from all studies for each of the outcomes of interest for the review question, when possible. **Quality assessment**: Evidence was critically appraised using the appropriate study design checklists as specified in Developing NICE guidelines: the manual. The overall quality of evidence by outcome was assessed using GRADE, which included 4 main quality elements (i.e., risk of bias, indirectness, inconsistency and imprecision). It was unclear if this process was conducted in duplicate.		indicator where variability in practice is expected **Guideline validation**: The draft guideline was posted for a 4-week stakeholder feedback period (2 weeks for healthcare professionals and patient support groups, 2 weeks for manufacturing companies). Comments received during this period were reviewed by the guideline development groups and recommendations were revised, if appropriate.
Ooi et al. (2019)[20] Representative countries included Malaysia, Thailand, Sri Lanka, India, China, Hong Kong, Taiwan, Philippines, Indonesia, Australia, New Zealand, Japan, South Korea, and Singapore. **Funding source**: Unrestricted educational grants from Medtronic, Takeda, LF Asia,	**Scope**: Best practices related to the use of immunomodulators and biologic agents for ulcerative colitis and Crohn's disease in Asia. **Developing institutions**: Asia Pacific Association of Gastroenterology Working Group on Inflammatory Bowel Disease and Asian Organization for Crohn's and Colitis. **Intended users**: Clinicians treating patients with ulcerative colitis or Crohn's disease. **Target population**: Asian populations (of all ages) with	**Evidence collection**: A systematic literature review was conducted in MEDLINE, Embase, and the Cochrane Trials Registry databases. In addition, regional and international consensus statements and guidelines on inflammatory bowel disease were examined. **Evidence selection**: No information on the methods of evidence selection were reported. **Evidence synthesis**: Evidence was categorized according to the Canadian	A modified Delphi process was used to develop the consensus statements according to their clinical importance. The steering committee generated a list of draft statements based on the available evidence. The list of statements was then presenting to the consensus group panel for discussion, revision, and voting. Consensus was achieved on a recommendation when 80% or more of votes were accepted "completely" or "with some reservation." Draft statements were	**Recommendation grading system**: Recommendations were classified as one of the following (according to the Canadian Task Force on the Periodic Health Examination): - A: There is good evidence to support the statement - B: There is fair evidence to support the statement - C: There is poor evidence to support the statement but recommendation made on other ground - D: There is fair evidence to refute the statement. - E: There is good evidence to refute the statement

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
the Journal of Gastroenterology and Hepatology Foundation, and the Asian Pacific Association of Gastroenterology.	ulcerative colitis or Crohn's disease.	Task Force on the Periodic Health Examination. No additional information on the processes for evidence synthesis were provided. **Quality assessment**: No information on the methods of quality assessment were reported.	refuted when 80% or more of voting members voted against the statement "completely" or "with some reservation."	**Guideline validation**: There was no mention of guideline validation
Rodríguez-Jiménez et al. (2018)[21] Spain **Funding source**: NR.	**Scope**: The prevention and treatment of TB infection in dermatology patients who are candidates for biologic therapy. **Developing institutions**: Spanish Academy of Dermatology and Venereology, Spanish Society of Pneumology and Thoracic Surgery, Spanish Society of Digestive Pathology, Spanish Society of Rheumatology, and Spanish Society of Infectious Diseases and Clinical Microbiology. **Intended users**: Dermatological specialists. **Target population**: Patients with dermatologic conditions who are candidates for biologic therapy.	**Evidence collection**: Literature searches were conducted in Medline and Cochrane databases up to July 2017. In addition, Spanish and international guidelines related to biological diseases and the diagnosis and treatment of latent TB infection were consulted. **Evidence selection**: No information on the methods of evidence selection were reported. **Evidence synthesis**: No information on the methods of evidence synthesis were reported. **Quality assessment**: It is unclear if quality assessment of identified literature was conducted.	Recommendations were drafted by participating experts using consensus. No additional information on how the recommendations were drafted and evaluated was provided.	**Recommendation grading system**: Recommendations were classified as one of the following: - A: Strong (recommendation supported) - B: Moderate (recommendation supported) - C: Weak (recommendation cannot be either supported or rejected) Recommendations were also given a grade based on their scientific quality: - Grade I: Recommendation based on at least 1 RCT with an appropriate design - Grade II: Recommendation based on at least 1 trial with an appropriate design but not randomized, cohort studies, multiple series, or very evident results in uncontrolled trials - Grade III: Recommendation based on expert opinion in descriptive studies or in clinical experience **Guideline validation**: There was no mention of guideline validation
GDG of the CPG on SLE (2016)[22] Spain	**Scope**: The diagnosis and management of SLE and its associated clinical manifestations.	**Evidence collection**: Systematic searches were conducted in Medline and PreMedline via OvidSP,	Recommendations were formulated by the development group following a consideration	**Recommendation grading system**: Recommendations were classified as one of the following (based on the SIGN criteria):

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
Funding source: The guideline was funded by the Institute of Health Carlos III and the Canary Islands Health service.	**Developing institutions**: Spanish Rheumatology Society, Spanish Society of Internal Medicine, Spanish Nephrology Society, Spanish Society of Haematology and Haemotherapy, Spanish Academy of Dermatology and Venereology, Spanish Neurology Society, Spanish Society of Primary Health Care Physicians, Spanish Hospital Pharmacy Society, Spanish Society of Primary Health Care Pharmacists, Spanish Society of Family and Community Medicine, Spanish Nursing Scientific Society. **Intended users**: Clinicians who have direct contact with SLE patients and who make clinical decisions relating to their care (e.g., rheumatologists, internists, nephrologists, dermatologists, haematologists, general practitioners, specialized nursing staff). **Target population**: Adults with SLE according to diagnostic criteria of expert physicians.	Embase via Elsevier and Science Citation Index Expanded and the Soial Science Citation Index via Web of Knowledge, The Cochrane Library, Psycinfo, Scopus, TripDatabase, Canadian Medical Association Infobase, International Guidelines Library, National Guidelines Clearinghouse, National Institute for Health and Care Excellence, SIGN, New Zealand Guidelines Group, Institute for Clinical Systems Improvement, and National Health and Medical Research Council between May and December 2014. Relevant study designs included guidelines, systematic reviews, RCTs, observational studies, and diagnostic test studies. **Evidence selection**: No information on the methods of evidence selection were reported. **Evidence synthesis**: Evidence was grouped using synthesis tables and summarized by clinical question. **Quality assessment**: The quality of identified studies was assessed using SIGN recommendations. Diagnostic studies ere assessed using the	for the evidence available and the equilibrium between desirable and undesirable consequences of carrying out the recommendations. Recommendations were drafted and agreed by consensus following face-to-face meetings by the development group and a series of consultation rounds with a panel of experts. A convergence of opinions was facilitated using a modified Delphi process.	- A: At least one MA, SR, or clinical trial classified as 1++ and directly appliable to the target population of the guidelines; or a volume of scientific evidence comprised of studies classified as 1+ and with great consistency between them. - B: A volume of scientific evidence comprised of studies classified as 2++, directly applicable to the target population of the guideline and that show great consistency between them; or scientific evidence extrapolated from studies classified as 1++ or 1+. - C: A volume of identified evidence comprised of studies classified as 2+, directly applicable to the target population of the guideline and that show great consistency between them; or scientific evidence extrapolated from studies classified as 2++. - D: Scientific evidence as level 3 or 3; or scientific evidence extrapolated from studies classified as 2+. - ✓: Recommended practice based on clinical experience and the consensus of the drafting team Scientific evidence for any given clinical question was classified as one of the following: were also given a grade based on their scientific quality: - 1++: High quality MA, SR of clinical trials, or high-quality clinical trials with very little bias risk

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
		Oxford Evidence-Based Medicine Centre system		- 1+: Well-performed MA, SR of clinical trials, or well-performed clinical trials with little bias risk - 1-: MA, SR of clinical trials, or clinical trials with high risk of bias - 2++: High-quality SR o case control or cohort studies. Well-conducted studies of case control or cohort studies with very low risk of confounding or bias and a high probability that the relationship is casual - 2+: Well-conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship Is causal - 2-: Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal - 3: Non-analytic studies, such as case reports and case series - 4: Expert opinion For questions related to diagnostic tests, recommendations were classified as follows: - A: Based on level Ia or Ib evidence - B: Based on level II evidence - C: Based on level III evidence - D: Based on level IV evidence Diagnostic evidence was classified into the following levels: - Ia: SR with homogenous level 1 studies - Ib: Level 1 studies - II: Level 2 studies; SR of level 2 studies - III: Level 3 studies; SR of level 3 studies

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation

				- IV: Consensus, expert opinions without explicit critical evaluation
				- Level 1 studies: The have a blinded comparison with a valid (gold standard) comparator test and have a suitable range of patients
				- Level 2 studies: Show one of these biases:
				  ○ Non-representative population
				  ○ Comparison with unsuitable comparator
				  ○ Non-blinded comparison
				  ○ Case and control studies
				- Level 3 studies: They meet two or more of the criteria for bias in level 2 studies
				
				Guideline validation: External reviewers examined the draft guidelines to improve the overall quality, to ensure appropriateness of recommendation, to disseminate the evidence, and to assess the applicability and feasibility of the recommendations.

Santin et al. (2016)[23] Spain **Funding source**: The guideline panel received no funding from for-profit organizations. The Spanish Society of Respiratory Diseases and Thoracic Surgery (SEPAR) and the	**Scope**: The guideline provides guidance regarding the use of IGRAs for diagnosing TB infection in immunocompetent and immunosuppressed adults and children of any age at risk for or suspected of having active TB. **Developing institutions**: Mycobacteria Study Group of the Spanish Society of Infectious Diseases and Clinical	**Evidence collection**: Literature searches were conducted in MEDLINE (accessed through PubMed) and Embase (accessed through Ovid), without language or temporal limitations, up to March 2013. Updates were performed until June 2015. The search was designed to identify relevant systematic reviews of diagnostic studies and additional relevant primary studies that updated existing	Recommendations were drafted after the panel considered the evidence available for each clinical question according to the GRADE methodology. The panel weighed the overall quality of the evidence, the balance between benefits and harms, the relative importance of the outcomes, and the resource use and costs. Recommendations were formalized after a	**Recommendation grading system**: Each recommendation was classified as “strong” or “weak”. **Guideline validation**: The guideline lists several experts who externally reviewed the document prior to its publication.

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
Spanish Society of Infectious Diseases and Clinical Microbiology (SEIMC) provided financial support for travel and housing costs to attend the in-person meetings.	Microbiology and the Spanish Society of Respiratory Diseases and Thoracic Surgery. **Intended users**: Primary health-care providers, specialists, and public health authorities who care for adults and children at risk for or suspect of having active tuberculosis infection. **Target population**: Immunocompetent and immunosuppressed adults and children of any age at risk for or suspected of having active TB.	systematic reviews. Studies that examined non-commercial or outdated IGRAs. Additional publications that provided information on resource use and costs were identified by searching in the NHS Economic Evaluation Database up to October 2014. **Evidence selection**: No information on the methods of evidence selection were reported. **Evidence synthesis**: Data extraction was performed in duplicate using a standardized data extraction sheet. Findings were summarized for each clinical question, which were formulated using a PICO format (i.e., population, intervention [index test], comparison [reference test] and outcomes). **Quality assessment**: The quality of evidence was assessed according to the GRADE criteria. This took into consideration study limitations, consistency between the results of the different studies, availability of direct evidence, precision of the estimators of effect, and publication bias. Following the evaluation of primary studies, the overall quality of evidence for each outcome was classified into	consensus between the members of the panel was reached.	

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
		four categories (i.e., high, moderate, low, and very low).		
Singh et al. (2015)[24] United States **Funding source**: ACR.	**Scope**: The guideline covers a wide-range of topics regarding the treatment and management of individuals with rheumatoid arthritis. **Developing institutions**: American College of Rheumatology. **Intended users**: Clinicians and their patients with rheumatoid arthritis. **Target population**: Adults (≥18 years) with rheumatoid arthritis.	**Evidence collection**: A series of systematic literature searches were conducted in OVID Medline, PubMed, Embase, and the Cochrane Library (including Cochrane Database of Systematic Reviews; Database of Abstracts of Reviews of Effects; Cochrane Central Register of Controlled Trials; and Health Technology Assessments) to identify literature that addressed the research questions. Relevant study designs included SRs, RCTs, and observational studies (including case series). For treatment modalities covered in past American College of Rheumatology guidelines, the searches were limited between January 1, 2009 and March 3, 2014. For treatment modalities not previously reviewed searches were conducted from inception up to March 3, 2014. Updates were performed on September 17, 2014. **Evidence selection**: Literature selection was performed in duplicate from a pool of reviewers. **Evidence synthesis**: The GRADE methodology was used for determining which literature to include for each	The voting panel formulated recommendations based on the relative benefits and harms of the treatment options under consideration, the quality of the underlying evidence, and patients' values and preferences (as per GRADE methodology). A consensus process was used.	**Recommendation grading system**: Recommendations were classified as being "strong" or "conditional", based on the underlying evidence that was ranked as being of high, moderate, low, or very low quality (using the GRADE system). - Strong recommendation: the panel was confident that the desirable effects of implementing the recommendation outweigh the undesirable effects (or vice versa). The recommendation would apply to most patients. - Conditional recommendation: the desirable effects of implementing the recommendation probably outweigh the undesirable effects. The recommendation would apply to the majority of patients, but some may not want to follow it. Conditional recommendations are preference sensitive and always warrant a shared decision-making approach. **Guideline validation:** In addition to the journal peer-review process, the guidelines were externally reviewed by the ACR Guideline Subcommittee, the ACR Quality of Care Committee, and the ACR Board of Directors.

Guideline citation, country, funding source	Scope, developing institutions, intended users, target population	Evidence collection, selection, synthesis, and quality assessment	Recommendations development and evaluation	Recommendation grading system, guideline validation
		research question. Data were pooled and analyzed in RevMan to determine an overall effect size for each research question. **Quality assessment**: Conducted in duplicate using the GRADE quality assessment criteria which considers the risk of bias in included trials, the likelihood of publication bias, inconsistency between trial results, indirectness of the evidence (e.g., heterogeneity), and imprecision.		

ACR = American College of Rheumatology; CPG = clinical practice guideline; GRADE = Grading of Recommendations Assessment, Development and Evaluation; GDG = Guideline Development Group; IGRA = interferon-gamma release assay; SLE = systemic lupus erythematosus; MA = meta-analysis; NICE = National Institute for Health and Care Excellence; NR = not reported; RCT = randomized controlled trial; SIGN = Scottish Intercollegiate Guidelines Network; SR = systematic review; TB = tuberculosis.

Appendix 3: Critical Appraisal of Included Publications

Table 5: Strengths and Limitations of the Systematic Review Using AMSTAR 2[10]

Strengths	Limitations
Alrajhi et al. (2020)[14]	
• The objectives and inclusion criteria were clearly stated and included components of population, intervention, comparator, and outcomes • The review methods were established prior to conducting the review (a protocol was prospectively registered on PROSPERO) • Multiple databases were searched (MEDLINE, Embase, and Cochrane Library databases). The electronic searches were supplemented with recursive searches, cross-referencing, and hand-searches • Key search terms (e.g., inflammatory bowel diseases, tuberculin test, interferon-gamma release tests) and publication restrictions were provided • Study selection, data extraction, and quality assessment processes were described and conducted in duplicate (disagreements were resolved through discussion and consensus, involving a third reviewer if required) • A flow chart of study selection was provided • The review authors described the included primary studies in adequate detail • The risk of bias of included primary studies was assessed using a satisfactory technique (i.e., the QUADAS-2) • Appropriate methods for the statistical combination of results were used in the meta-analyses • Statistical heterogeneity was assessed using the I^2 statistic • Review authors stated that their potential conflicts of interest (one author was a speaker or an advisory board member for Janssen, AbbVie, Takeda, Pfizer, Shire, Ferring, PendoPharm, and Merck and has received unrestricted research grants from Janssen, Abbvie, Pentax, and Echosense; all authors had no conflicts of interests to declare pertaining to this publication) • Source of funding was disclosed (there was no funding received for this review)	• It was unclear which study designs were eligible for inclusion and the authors did not explain their selection of study designs for inclusion in the review • A grey literature search was not completed • A list of excluded studies was not provided (although the reasons for exclusion were) • Review authors did not report on sources of funding for the included primary studies • The risk of bias and limitations of primary study methodology were not adequately considered when discussing the results • There was no discussion on the possibility of publication bias

AMSTAR 2 = A MeaSurement Tool to Assess systematic Reviews 2; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2.

Table 6: Strengths and Limitations of the Non-Randomized Study Using the Downs and Black checklist[11]

Strengths	Limitations
Chaer et al. (2020)[15]	
• The objectives, interventions, and main outcomes were clearly described • Patient eligibility criteria was provided • Participant characteristics (e.g., age, gender, diagnosis, duration of disease) were clearly described • The main findings of the study were clearly described • Due to the nature of the study (i.e., a retrospective descriptive observational study), no participants were lost to follow-up • Study participants, care providers, and setting appeared to be representative of the population and care setting of interest • Compliance with the intervention was reliable • The authors declared that they had no potential conflicts of interest • Sources of funding were disclosed (UCB Brazil) and were unlikely to have had an effect on the findings of the study	• The study had no comparator (i.e., it did not include a control group); therefore, the results are susceptible to numerous forms of bias and any findings should be considered associative (rather than causative). For example, without a control group we cannot be certain that the use of the TST prior to initiating biologic therapy had any impact on the number of participants who developed active TB • This is a descriptive study that aimed to describe the clinical outcomes (e.g., incidence of active TB infection) of a cohort of patients undergoing biologic therapy. Statistical tests were not conducted to estimate the risk of developing TB after initiating biologic therapy • Adverse events that may have been associated with the use of the TST were not reported • This was an open-label study with no blinding of study participants or outcome assessors • Single-centre study (conducted in the Brazil); the generalizability to the Canadian setting was unclear

TB = tuberculosis; TST = tuberculin skin test.

Table 7: Strengths and Limitations of Diagnostic Test Accuracy Studies Using the QUADAS-2 Checklist[12]

Strengths	Limitations
Girit et al. (2020)[16]	
• The objectives, interventions, controls, outcomes, and main findings were clearly described • While not explicitly stated, participants appear to have been recruited consecutively • A case-control study design was avoided • Patient inclusion and exclusion criteria were provided • Inappropriate exclusion criteria were avoided • The cut-off value for a positive TST was pre-specified • Population characteristics (e.g., age, sex, clinical diagnosis, time since diagnosis, current and past medical therapy) were clearly described • There was an appropriate time interval between the two diagnostic tests (i.e., they were given simultaneously) • All participants were evaluated with the same diagnostic tests • Study participants, care providers, and setting appear to be representative of the population and care setting of interest • The authors declared that they had no potential conflicts of interest • Source of funding was disclosed (there was no funding received for this study)	• Because there was no gold standard for detecting latent TB infection, Cohen's kappa was used to evaluate the coherence and inter-rater agreement between the TST and IGRA • It was unclear if the two diagnostic tests were interpreted without knowledge of the result of the other test • In some of their analyses, diagnostic parameters were calculated using TST as the reference standard; however, the TST is not considered a gold standard and is not likely to have correctly classified the target condition • This study was conducted in four tertiary Pediatric Pulmonology Centers in Turkey; the generalizability to the Canadian setting is unclear

Strengths	Limitations
Sellami et al. (2019)[17]	
• The objectives, interventions, controls, outcomes, and main findings were clearly described • While not explicitly stated, participants appear to have been recruited consecutively • A case-control study design was avoided • Patient inclusion and exclusion criteria were provided • Inappropriate exclusion criteria were avoided • The two diagnostic tests were interpreted by clinicians who were blinded to the result of the other test • The cut-off value for a positive TST was pre-specified • Population characteristics (e.g., age, sex, clinical diagnosis, time since diagnosis, current and past medical therapy) were clearly described • There was an appropriate time interval between the two diagnostic tests (i.e., they were given on the same day) • All participants were evaluated with IGRA • Study participants, care providers, and setting appear to be representative of the population and care setting of interest • The authors declared that they had no potential conflicts of interest • Source of funding was disclosed (there was no funding received for this study)	• Diagnostic parameters were calculated for both the TST and IGRA using chest CT as the reference standard; however, chest CT is not considered a gold standard and will not always correctly identify latent TB • TST was not performed on all participants (10.5% of participants did not received TST) • This study was conducted in a single hospital in Tunisia; the generalizability to the Canadian setting is unclear
So et al. (2017)[18]	
• The objectives, interventions, outcomes, and main findings were clearly described • While not explicitly stated, participants appear to have been recruited consecutively • A case-control study design was avoided • Patient inclusion and exclusion criteria were provided • Inappropriate exclusion criteria were avoided • The two diagnostic tests were interpreted by clinicians who were blinded to the result of the other test • The cut-off value for a positive TST was pre-specified • Population characteristics (e.g., age, sex, clinical diagnosis, time since diagnosis, current and past medical therapy) were clearly described • There was an appropriate time interval between the two diagnostic tests (i.e., they were given simultaneously) • All participants were evaluated with the same diagnostic tests • Study participants, care providers, and setting appear to be representative of the population and care setting of interest • The authors declared that they had no potential conflicts of interest	• Because there was no gold standard for detecting latent TB infection, both the TST and IGRA were considered index tests and Cohen's kappa was used to evaluate the coherence and inter-rater agreement between the two tests • The study's source of funding was not reported • This study was conducted in a single hospital in Hong Kong; the generalizability to the Canadian setting is unclear

CT = computerized tomography; IGRA = interferon-gamma release assay; QUADAS-2 = Quality Assessment of Diagnostic Accuracy Studies 2; TB = tuberculosis; TST = tuberculin skin test.

Table 8: Strengths and Limitations of Guidelines Using AGREE II[13]

Item	Guideline					
	Smith et al. (2020)[19]	Ooi et al. (2019)[20]	Rodríguez-Jiménez et al. (2018)[21]	GDG of the CPG on SLE (2016)[22]	Santin et al. (2016)[23]	Singh et al. (2015)[24]
Domain 1: Scope and Purpose						
1. The overall objective(s) of the guideline is (are) specifically described.	Yes	Yes	Yes	Yes	Yes	Yes
2. The health question(s) covered by the guideline is (are) specifically described.	Yes	Yes	Yes	Yes	Yes	Yes
3. The population (patients, public, etc.) to whom the guideline is meant to apply is specifically described.	Yes	Yes	Yes	Yes	Yes	Yes
Domain 2: Stakeholder Involvement						
4. The guideline development group includes individuals from all relevant professional groups.	Yes	Yes	Yes	Yes	Yes	Yes
5. The views and preferences of the target population (patients, public, etc.) have been sought.	Yes	No	No	Yes	No	Yes
6. The target users of the guideline are clearly defined.	Yes	Yes	Yes	Yes	Yes	Yes
Domain 3: Rigour of Development						
7. Systematic methods were used to search for evidence.	Yes	Partially	Partially	Yes	Yes	Yes
8. The criteria for selecting the evidence are clearly described.	Yes	No	No	Partially	Partially	Yes
9. The strengths and limitations of the body of evidence are clearly described.	Yes	Partially	Yes	Yes	Yes	Yes
10. The methods for formulating the recommendations are clearly described.	Yes	Yes	Partially	Yes	Partially	Yes
11. The health benefits, side effects, and risks have been considered in formulating the recommendations.	Yes	Yes	Yes	Yes	Yes	Yes

Item	Guideline					
	Smith et al. (2020)[19]	Ooi et al. (2019)[20]	Rodríguez-Jiménez et al. (2018)[21]	GDG of the CPG on SLE (2016)[22]	Santin et al. (2016)[23]	Singh et al. (2015)[24]
12. There is an explicit link between the recommendations and the supporting evidence.	Yes	Yes	Yes	Yes	Yes	Yes
13. The guideline has been externally reviewed by experts prior to its publication.	Yes	No	No	Yes	Yes	Yes
14. A procedure for updating the guideline is provided.	Yes	No	No	Yes	Yes	Yes
Domain 4: Clarity of Presentation						
15. The recommendations are specific and unambiguous.	Yes	Yes	Yes	Yes	Yes	Yes
16. The different options for management of the condition or health issue are clearly presented.	Yes	Yes	Yes	Yes	Yes	Yes
17. Key recommendations are easily identifiable.	Yes	Yes	Yes	Yes	Yes	Yes
Domain 5: Applicability						
18. The guideline describes facilitators and barriers to its application.	Yes	No	No	Yes	Yes	No
19. The guideline provides advice and/or tools on how the recommendations can be put into practice.	Yes	Yes	Yes	Yes	Yes	Yes
20. The potential resource implications of applying the recommendations have been considered.	Yes	Yes	Yes	Yes	Yes	Yes
21. The guideline presents monitoring and/or auditing criteria.	Yes	No	No	Yes	No	No
Domain 6: Editorial Independence						
22. The views of the funding body have not influenced the content of the guideline.	Yes	Yes	Unclear	Yes	Yes	Yes

Item	Guideline					
	Smith et al. (2020)[19]	Ooi et al. (2019)[20]	Rodríguez-Jiménez et al. (2018)[21]	GDG of the CPG on SLE (2016)[22]	Santin et al. (2016)[23]	Singh et al. (2015)[24]
23. Competing interests of guideline development group members have been recorded and addressed.	Yes	Yes	Yes	Yes	Yes	Yes

AGREE II = Appraisal of Guidelines for Research and Evaluation II; CPG = clinical practice guideline; GDG = Guideline Development Group; SLE = systemic lupus erythematosus.

Appendix 4: Main Study Findings and Authors' Conclusions

Table 9: Summary of Findings of the Included Systematic Review

Main study findings	Authors' conclusion
Alrajhi et al. (2020)[14]	

Systematic review with meta-analyses that assessed test concordance between TSTs and IGRAs for latent TB screening in patients with inflammatory bowel disease prior to initiating anti-TNF medications. A total of 16 studies were included in the review.
 	
Summary of findings
- The review authors combined data from all 16 included primary studies (N = 2,488) and reported a pooled concordance between TST and IGRA of 85% (95% CI = 81% to 88%; P value = 0.01). There was notable heterogeneity in this result (I^2 – 76%; P value = 0.001).
- Additional meta-analyses were conducted in a subset of eight studies (N = 811) to determine the odds ratio for testing positive by IGRA and TST if participants were immunosuppressed
  ○ Using a random effects model, the odds ratio of testing positive by IGRA in immunosuppressed individuals was 0.57 (95% CI = 0.31 to 1.03; P value = 0.06). This result had possible no important heterogeneity (I^2 = 32%; P = 0.17).
  ○ Using a fixed effects model, the odds ratio of testing positive by IGRA in immunosuppressed individuals was 0.57 (95% CI = 0.37 to 0.89; P value = 0.01). This result had possible no important heterogeneity (I^2 = 32%; P = 0.17).
  ○ The odds ratio of testing positive by TST in immunosuppressed individuals was 1.14 (95% CI = 0.61 to 2.12; P value = 0.69). This result had possible moderate heterogeneity (I^2 = 42%; P = 0.10).
	"While concordance was 85% between TST and IGRA, the performance of IGRA seems to be negatively affected by immunosuppression. Given the importance of detecting latent TB prior to anti-TNF initiation, further randomized controlled trials comparing the performance of TST and IGRA in IBD patients are needed (p. 306)."[14]

CI = confidence interval; IGRA = interferon-gamma release assay; N = number of participants; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Table 10: Summary of Findings of Included Primary Clinical Studies

Main study findings	Authors' conclusion
Non-randomized study	
Chaer et al. (2020)[15]	
Single centre observational study that evaluated the frequency of TB infection in patients with rheumatoid arthritis, ankylosing spondylitis, juvenile idiopathic arthritis, and psoriatic arthritis (N = 161) after the use of biologic agents. Summary of findings - Of the 161 rheumatic patients included in the study, 31 (19.25%) were diagnosed with latent TB infection according to the TST prior to starting biologic therapy. All positive patients received latent TB prophylaxis with isoniazid (300 mg daily) for six months. Biologic treatment was started only after the patient received isoniazid for at least two months. - Throughout the course of the study, 11 of the 161 patients (6.8%) developed active TB (four of these individuals tested positive for the TST and developed active TB despite prophylactic treatment) ○ Among those who developed active TB, six had ankylosing spondylitis (54.5%), one had ankylosing spondylitis (9.09%), two had rheumatoid arthritis (18.18%), and two had juvenile idiopathic arthritis (18.18%)	"In this study, the frequency of TB infection among 161 consecutive patients on biologics, during a follow-up of 55 months, was 6.8%. Despite some limitations, this study reinforces that in a TB [high-burden country], the possibility of TB infection in a patient receiving biological therapy should always be considered even after the LTBI treatment with isoniazid. Compared with national data from BiobadaBrasil, we found a higher incidence of TB (6.8

Main study findings	Authors' conclusion
	versus 0.44%). Multicenter and long-term studies could provide more data about this important health issue (p. 5)."[15]
Diagnostic test accuracy studies	
Girit et al. (2020)[16]	

Multi-centre, prospective diagnostic test accuracy study that investigated the coherence between IGRA and TST with different cut-off values for screening latent TB infection prior to and after six months of treatment with anti-TNF therapy in children (<18 years of age) who were diagnosed with immune-mediated inflammatory diseases (N = 57).
 	
<u>Summary of findings</u>
- **Prior to treatment with biologic therapy (using a ≥5 mm TST cut-off)**
  ○ Positive result:
  ▪ IGRA: N = 5 (8.8%)
  ▪ TST: N = 16 (28.1%)
  ○ Negative result:
  ▪ IGRA: N = 52 (91.2%)
  ▪ TST: N = 41 (71.9%)
  ○ Diagnostic parameters
  ▪ Cohen's kappa (κ): 0.176
  ▪ Strength of agreement: Poor
  ▪ Sensitivity: 60.00%
  ▪ Specificity: 75.00%
  ▪ Positive predictive value: 18.75%
  ▪ Negative predictive value: 95.12%
  ▪ Accuracy: 73.68%
- **Prior to treatment with biologic therapy (using a 5-9 mm TST cut-off)**
  ○ Positive result:
  ▪ IGRA: N = 3 (6.4%)
  ▪ TST: N = 6 (12.8%)
  ○ Negative result:
  ▪ IGRA: N = 44 (93.6%)
  ▪ TST: N = 41 (87.2%)
  ○ Diagnostic parameters
  ▪ Cohen's kappa (κ): 0.150
  ▪ Strength of agreement: Poor
  ▪ Sensitivity: 33.33%
  ▪ Specificity: 88.64%
  ▪ Positive predictive value: 16.67%
  ▪ Negative predictive value: 95.12%
  ▪ Accuracy: 85.11%
- **Prior to treatment with biologic therapy (using a 10-14 mm TST cut-off)**
  ○ Positive result:
  ▪ IGRA: N = 4 (7.5%)
  ▪ TST: N = 6 (11.3%)
  ○ Negative result:
  ▪ IGRA: N = 49 (92.5%)
  ▪ TST: N = 47 (88.7%)
  ○ Diagnostic parameters
  ▪ Cohen's kappa (κ): 0.120
  ▪ Strength of agreement: Poor
	"In conclusion, each country should develop its own guidelines according to the BCG status and TB disease prevalence for LTBI screening prior to anti-TNF treatment. Our results show that in our country, with all patients vaccinated, there is poor coherence between TST and [IGRA] in screening both prior to and under anti-TNF treatment. Under these circumstances, it would still be rational to use both tests for screening until further studies bring new evidence on the subject (p. 26)."[16]

Main study findings	Authors' conclusion
▪ Sensitivity: 25.00% ▪ Specificity: 89.80% ▪ Positive predictive value: 16.67% ▪ Negative predictive value: 93.62% ▪ Accuracy: 84.91% - **Prior to treatment with biologic therapy (using a ≥15 mm TST cut-off)** ○ Positive result: ▪ IGRA: N = 5 (8.8%) ▪ TST: N = 4 (7.0%) ○ Negative result: ▪ IGRA: N = 52 (91.2%) ▪ TST: N = 53 (93.0%) ○ Diagnostic parameters ▪ Cohen's kappa (κ): 0.156 ▪ Strength of agreement: Poor ▪ Sensitivity: 20.00% ▪ Specificity: 94.23% ▪ Positive predictive value: 25.00% ▪ Negative predictive value: 92.45% ▪ Accuracy: 87.72% - **After six months of treatment with biologic therapy (using a ≥5 mm TST cut-off)** ○ Positive result: ▪ IGRA: N = 5 (8.8%) ▪ TST: N = 19 (33.3%) ○ Negative result: ▪ IGRA: N = 52 (91.2%) ▪ TST: N = 38 (66.7%) ○ Diagnostic parameters ▪ Cohen's kappa (κ): 0.129 ▪ Strength of agreement: Poor ▪ Sensitivity: 60.00% ▪ Specificity: 69.23% ▪ Positive predictive value: 15.79% ▪ Negative predictive value: 94.74% ▪ Accuracy: 68.42% - **After six months of treatment with biologic therapy (using a 5-9 mm TST cut-off)** ○ Positive result: ▪ IGRA: N = 2 (4.3%) ▪ TST: N = 8 (17.4%) ○ Negative result: ▪ IGRA: N = 44 (95.7%) ▪ TST: N = 38 (82.6%) ○ Diagnostic parameters ▪ Cohen's kappa (κ): - ▪ Strength of agreement: No agreement ▪ Sensitivity: 0% ▪ Specificity: 81.82% ▪ Positive predictive value: 0% ▪ Negative predictive value: 94.74% ▪ Accuracy: 78.26% - **After six months of treatment with biologic therapy (using a 10-14 mm TST cut-off)** ○ Positive result: ▪ IGRA: N = 4 (7.7%) ▪ TST: N = 6 (11.5%) ○ Negative result:	

Main study findings	Authors' conclusion
▪ IGRA: N = 48 (92.3%) ▪ TST: N = 46 (88.5%) ○ Diagnostic parameters ▪ Cohen's kappa (κ): 0.339 ▪ Strength of agreement: Fair ▪ Sensitivity: 50.00% ▪ Specificity: 91.67% ▪ Positive predictive value: 33.33% ▪ Negative predictive value: 95.65% ▪ Accuracy: 88.46% - **After six months of treatment with biologic therapy (using a ≥15 mm TST cut-off)** ○ Positive result: ▪ IGRA: N = 5 (8.8%) ▪ TST: N = 5 (8.8%) ○ Negative result: ▪ IGRA: N = 52 (91.2%) ▪ TST: N = 52 (91.2%) ○ Diagnostic parameters ▪ Cohen's kappa (κ): 0.123 ▪ Strength of agreement: Poor ▪ Sensitivity: 20.00% ▪ Specificity: 92.00% ▪ Positive predictive value: 20.00% ▪ Negative predictive value: 92.00% ▪ Accuracy: 85.45% - In addition to these findings, the authors noted that "There was no statistically significant difference in TST positivity rates between the patient groups with or without glucocorticoid use (for both low dose and high dose glucocorticoid use), patient groups using different [disease-modifying antirheumatic drugs] and patient groups who were planned to use different anti-TNFs"[16] (p. 22). Numerical data for these findings was not presented.	
Sellami et al. (2019)[17]	
Single-centre, retrospective diagnostic test accuracy study that evaluated the agreement between IGRA and the TST in detecting latent TB infection in patients with chronic immune-mediated inflammatory diseases prior to biologic therapy (N = 105). <u>Summary of findings</u> - **Results of the diagnostic tests** ○ All participants (N = 105; using a 10 mm TST cut-off): ▪ TST+ and IGRA+: N = 18 (17.1%) ▪ TST- and IGRA-: N = 45 (42.8%) ▪ TST+ and IGRA-: N = 25 (23.8%) ▪ TST- and IGRA+: N = 5 (4.7%) ▪ TST not tested and IGRA+: N = 0 (0%) ▪ TST not tested and IGRA-: N = 11 (10.0%) - **Agreement between the two diagnostic tests** ○ All participants (N = 95; using a 10 mm TST cut-off): ▪ Cohen's kappa (κ): 0.08 (95% CI = 0.003 to 0.162) ▪ Strength of agreement: Low - **Diagnostic parameters** ○ Sensitivity (using chest CT as the reference standard) ▪ IGRA: 26.14% ▪ TST (10 mm cut-off): 73.9% ▪ TST (5 to 9 mm): 73.9% ▪ TST (10 to 14 mm): 73.9%	"In conclusion, the benefits and limitations of both TST and IGRA and the low agreement between the two tests have been confirmed. This makes the decision of choosing the gold standard for LTBI screening challenging and depending on the countries' incidence of TB and economic status. Given the most of the specificity, IGRAs should be included in the strategy to diagnose LTBI in patients with chronic inflammatory diseases before starting or switching biologic treatment, particularly as the substitution of TST by IGRA led to a significant reduction in the need of

Main study findings	Authors' conclusion
▪ TST (15 to 19 mm): 52.1% ▪ TST (20 to 24 mm): 39.1% ▪ TST (≥ 25 mm): 4.3% ○ Specificity (using chest CT as the reference standard) ▪ IGRA: 100% ▪ TST (10 mm cut-off): 80.6% ▪ TST (5 to 9 mm): 79.1% ▪ TST (10 to 14 mm): 80.6% ▪ TST (15 to 19 mm): 97.0% ▪ TST (20 to 24 mm): 98.5% ▪ TST (≥ 25 mm): 98.5% ○ Accuracy (using chest CT as the reference standard) ▪ IGRA: 75.3% ▪ TST (10 mm cut-off): 73.6% ▪ TST (5 to 9 mm): 71.2% ▪ TST (10 to 14 mm): 72.5% ▪ TST (15 to 19 mm): 74.1% ▪ TST (20 to 24 mm): 74.4% ▪ TST (≥ 25 mm): 75.6%	antibiotics prophylaxis (p. 229)."[17]
So et al. (2017)[18]	
Single-centre, retrospective diagnostic test accuracy study that evaluated the concordance between IGRA and TST in the diagnosis of latent TB in patients with arthritic disease prior to treatment with biologic therapies (N = 38). <u>Summary of findings</u> - **Results of the diagnostic tests** ○ All participants (N = 38; using a 10 mm TST cut-off): ▪ TST+ and IGRA+: N = 7 (18.4%) ▪ TST- and IGRA-: N = 21 (55.3%) ▪ TST+ and IGRA-: N = 6 (15.8%) ▪ TST- and IGRA+: N = 4 (10.5%) ○ Participants on prednisolone (N = 19; using a 10 mm TST cut-off): ▪ TST+ and IGRA+: N = 1 (5.3%) ▪ TST- and IGRA-: N = 12 (63.2%) ▪ TST+ and IGRA-: N = 4 (21.1%) ▪ TST- and IGRA+: N = 2 (10.5%) ○ Participants not on prednisolone (N = 19; using a 10 mm TST cut-off): ▪ TST+ and IGRA+: N = 6 (31.6%) ▪ TST- and IGRA-: N = 9 (47.4%) ▪ TST+ and IGRA-: N = 2 (10.5%) ▪ TST- and IGRA+: N = 2 (10.5%) ○ All participants (N = 38; using a 5 mm TST cut-off): ▪ TST+ and IGRA+: N = 10 (26.3%) ▪ TST- and IGRA-: N = 18 (47.4%) ▪ TST+ and IGRA-: N = 9 (23.7%) ▪ TST- and IGRA+: N = 1 (2.6%) ○ Participants not on prednisolone (N = 19; using a 5 mm TST cut-off): ▪ TST+ and IGRA+: N = 8 (42.1%) ▪ TST- and IGRA-: N = 8 (42.1%) ▪ TST+ and IGRA-: N = 3 (15.8%) ▪ TST- and IGRA+: N = 0 (0.0%) - **Agreement between the two diagnostic tests** ○ All participants (N = 38; using a 10 mm TST cut-off): ▪ Cohen's kappa (κ): 0.39 (SE = 0.16)	"In arthritis patients in a TB-endemic population, the level of agreement between TST and A.TB IGRA for detecting LTBI was only fair. Although 10 mm is the usual cut-off for TST, the level of agreement between the two tests improved from fair to moderate when a 5-mm cut-off was used. A dual testing strategy with TST and IGRA appeared to be effective and should be pursued, especially in patients who are prescribed a systemic steroid. The issue of potential overtreatment is yet to be evaluated (p. 250)."[18]

Main study findings	Authors' conclusion
▪ Strength of agreement: Fair ○ All participants (N = 38; using a 5 mm TST cut-off): ▪ Cohen's kappa (κ): 0.47 (SE = 0.13) ▪ Strength of agreement: Moderate ○ Participants on prednisolone (N = 19; using a 10 mm TST cut-off): ▪ Cohen's kappa (κ): 0.066 (SE = 0.23) ▪ Strength of agreement: Slight ○ Participants not on prednisolone (N = 19; using a 10 mm TST cut-off): ▪ Cohen's kappa (κ): 0.57 (SE = 0.16) ▪ Strength of agreement: Moderate ○ Participants not on prednisolone (N = 19; using a 5 mm TST cut-off): ▪ Cohen's kappa (κ): 0.69 (SE = 0.16) ▪ Strength of agreement: Substantial - In addition to these results, the authors noted that 32 of the 38 study participants ended up receiving biologic therapy. None of these participants developed active TB by the end of the one year follow-up.	

BCG = Bacillus Calmette–Guérin; CI = confidence interval; CT = computerized tomography; IGRA = interferon-gamma release assay; LTBI = latent tuberculosis infection; N = number of participants; SE = standard error; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Table 11: Summary of Recommendations in Included Guidelines

Recommendations and supporting evidence	Quality of evidence and strength of recommendations
Smith et al. (2020)[19]	
Evidence-based guidelines regarding the use of biologic therapies in individuals (of all ages) with psoriasis. **Relevant recommendations**: 1. **"**Consider screening for latent tuberculosis with an IGRA alone, or with an IGRA and concurrent Mantoux test; be aware of the individual's risk factors for TB when interpreting results (p. 634)."[19,20] ○ **Supporting evidence:** the guideline development group considered NICE guideline recommendations on diagnosing latent TB in adults	- **Quality of evidence**: NR - **Strength of recommendation**: ↑ (weak recommendation for the use of an intervention)
Ooi et al. (2019)[20]	
Evidence-based guideline regarding best practices related to the use of immunomodulators and biologic agents for ulcerative colitis and Crohn's disease in Asia. **Relevant recommendations**: 1. "Routine screening for LTBI should be performed according to local practice before initiating biologic treatment. This may include chest X-ray, chest computed tomography, IGRA and/or TST (p. 302)."[20] ○ **Supporting evidence:** six studies	- **Quality of evidence**: 2A (evidence obtained from well-designed cohort or case-control) - **Classification of the recommendation**: A (there is good evidence to support the statement)
Rodríguez-Jiménez et al. (2018)[21]	

Recommendations and supporting evidence	Quality of evidence and strength of recommendations
Evidence-based guideline regarding the prevention and treatment of TB infection in dermatology patients who are candidates for biologic therapy. **Relevant recommendations**:	
1. "All candidates for biological therapies should be studied for diagnosis of [LTBI] as these patients belong to one of the highest risk groups for developing TB (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
2. "The risk of these patients falling ill appears to be depend on the anti-TNF agent used: infliximab and adalimumab are those with highest risk detected (p. 595)."[21] ○ **Supporting evidence**: two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
3. "The diagnostic methods for [LTBI] detection are based on: - Active search for TB history in the patient's medical record as well as contacts with active TB - Chest X-ray can detect possible old TB lesions. In the case of doubt, it is advised to complete the study with chest CT, given that this technique is superior to chest X-ray for detection of early radiologic signs of active TB or old lesions - Simultaneous performance of IGRA and [TST]. Any positive result in one of these tests is considered indicative of [LTBI] (p. 595)."[21] ○ **Supporting evidence:** one study	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
4. "Patients diagnosed with immune-mediated inflammatory diseases have a higher rate of false negatives with [TST] and IGRA (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
5. "There is no evidence that repeating [TST] (booster effect) increases the sensitivity of the test in patients with immune-mediated inflammatory diseases but that it does reduce specificity, and so currently, with the use of IGRA, the [TST] is not recommended (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: C - Weak (recommendation cannot be either supported or rejected supported)
6. "IGRA should be performed before [TST], due to the booster effect detected for IGRAs (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
7. "The specificity and sensitivity of the 2 IGRA techniques for the diagnosis of [LTBI] are similar in patients with [immune-mediated inflammatory diseases]; as the sensitivity of *T.SPOT.TB* is somewhat higher in patients treated with corticosteroids, use of this test should be considered in these patients (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: B - Moderate (recommendation supported)

Recommendations and supporting evidence	Quality of evidence and strength of recommendations
8. "Indeterminate results should always be confirmed with a second measurement that, in most cases, is negative (p. 595)."[21] ○ **Supporting evidence:** two studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
9. "Negativity in the [TST] and IGRA does not rule out [LTBI] (p. 595)."[21] ○ **Supporting evidence:** four studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
10. **"**Preventive treatment is recommended in all candidates for therapy who presented with suspicion of LTI, after excluding active TB, in any of the diagnostic tests (p. 595)."[21] ○ **Supporting evidence:** five studies	- **Quality of evidence**: III (recommendation based on expert opinion in descriptive studies or in clinical experience) - **Strength of recommendation**: A - Strong (recommendation supported)
GDG of the CPG on SLE (2016)[22]	
Evidence-based guideline regarding the diagnosis and management of SLE and its associated clinical manifestations. **Relevant recommendations**:	
1. "We cannot give a general recommendation on the indication or periodicity of repeated assessments of latent infection due to the human immunodeficiency virus, the hepatitis B virus, the hepatitis C virus and tuberculosis. Therefore these should be adapted to the clinical situation and the individual risk factors of each patient (p. 45)."[22] ○ **Supporting evidence**: no data from clinical studies; however, a previously published consensus guideline was considered	- **Quality of evidence**: NA - **Strength of recommendation**: Recommended practice based on clinical experience and the consensus of the drafting team
2. "We suggest examining all patients who are going to be submitted to immunosuppressive treatment for human immunodeficiency virus, the hepatitis B virus, hepatitis C virus and tuberculosis, above all when this treatment involves high doses of glucocorticoids or biological therapies, regardless of the existence of risk factors (p. 45)."[22] ○ **Supporting evidence**: no data from clinical studies; however, a previously published consensus guideline was considered	- **Quality of evidence**: NA - **Strength of recommendation**: Recommended practice based on clinical experience and the consensus of the drafting team
3. "For patients whose first tuberculin skin test is negative, we suggest carrying out a second test one week later to induce the immunological memory (booster effect) as false negatives are more frequent in the elderly and in immunosuppressed patients (p. 45)."[22] ○ **Supporting evidence**: 1 non-randomized clinical trial	- **Quality of evidence**: based on scientific evidence of level 3; or scientific evidence extrapolated from studies classified as 2+ - **Strength of recommendation**: D
4. "The tuberculin skin test is the test of choice to detect tuberculosis thanks to its sensitivity in diagnosing tuberculosis in the standard cut-off point (5 mm). However, previous BCG vaccination and/or immunosuppression, could make the [QuantiFERON Gold Test] a more reliable test for detecting latent infection (p. 45)."[22]	- **Quality of evidence**: NA - **Strength of recommendation**: (recommended practice based on clinical experience and the consensus of the drafting team)

Recommendations and supporting evidence	Quality of evidence and strength of recommendations
○ **Supporting evidence**: no data from clinical studies; however, a previously published consensus guideline was considered	
Santin et al. (2016)[23]	
Evidence-based guideline regarding the use of IGRAs in the diagnosis of tuberculosis infection. **Relevant recommendations**: 1. "The panel suggests using both the TST and an IGRA to screen for TB infection in patients with chronic inflammatory disease before starting biological therapy (p. 303.e4)."[23] ○ **Supporting evidence**: Eight observational studies assessed the predictive values of IGRAs for the development of TB in patients treated with biologic therapies; seven studies assessed the predictive value of the TST for the development of TB; five studies assessed the impact of immunosuppression on the results of IGRAs and the TST.	- **Quality of evidence**: Low to Very low - **Strength of recommendation**: Weak
Singh et al. (2015)[24]	
Evidence-based guideline regarding the treatment and management of adults with rheumatoid arthritis. **Relevant recommendations**: 1. The voting panel endorsed a tuberculosis screening algorithm for patients who are candidates for biologics or tofacitinib. The algorithm indicated that initial screening should be performed with TST or IGRA. IGRA was preferred in patients with a history of BCG vaccination; anergy panel testing was not recommended. Rescreening was recommended if tuberculosis risk factors are present. If initial screening is negative the patient should proceed to start or resume biologic or tofacitinib immediately. Those who receive a positive result from their initial screen should be screened for active tuberculosis with a chest radiograph. If the chest radiograph result is negative patients should be given treatment for latent tuberculosis prior to starting their biologic therapy or tofacitinib. Positive radiograph should be confirmed with sputum for acid-fast bacilli. If the sputum test is negative treat for latent tuberculosis, if the sputum test is positive complete treatment for active tuberculosis prior to starting biologic therapy or tofacitinib. ○ **Supporting evidence**: NR	- **Quality of evidence**: NR - **Strength of recommendation**: NR

BCG = Bacillus Calmette–Guérin; CPG = clinical practice guideline; GDG = Guideline Development Group; GRADE = Grading of Recommendations Assessment, Development and Evaluation; IGRA = interferon-gamma release assay; LTBI = latent tuberculosis infection; NA = not applicable; NR = not reported; SLE = systemic lupus erythematosus; TB = tuberculosis; TNF = tumour necrosis factor; TST = tuberculin skin test.

Appendix 5: Further Information

Previous CADTH Reports

CADTH. Tuberculosis in People with Compromised Immunity: A Review of Guidelines; 2020. https://www.cadth.ca/tuberculosis-people-compromised-immunity-review-guidelines

Review Articles

Zellweger JP, Sotgiu G, Corradi M, Durando P. The diagnosis of latent tuberculosis infection (LTBI): currently available tests, future developments, and perspectives to eliminate tuberculosis (TB). *Med Lav*. 2020 Jun 26;111(3):170-183. PubMed: PM32624559

Fernandez-Ruiz M. Assessment of latent infections in patients receiving biological therapies. *Rev Esp Quimioter*. 2019 Sep;32 Suppl 2:63-68. PubMed: PM31475814