Comparing different modalities for the diagnosis of incisional hernia: a systematic review

L. F. Kroese1 · D. Sneiders1 · G. J. Kleinrensink2 · F. Muysoms3 · J. F. Lange1

Received: 10 November 2017 / Accepted: 27 December 2017 / Published online: 11 January 2018
© The Author(s) 2018. This article is an open access publication

Abstract
Purpose Incisional hernia (IH) is the most frequent complication after abdominal surgery. The diagnostic modality, observer, definition, and diagnostic protocol used for the diagnosis of IH potentially influence the reported prevalence. The objective of this systematic review is to evaluate the diagnostic accuracy of different modalities used to identify IH.

Methods Embase, MEDLINE OvidSP, Web of Science, Google Scholar, and Cochrane databases were searched to identify studies diagnosing IH. Studies comparing the IH detection rate of two different diagnostic modalities or inter-observer variability of one modality were included. Quality assessment of studies was done by Cochrane Collaboration’s tool. Article selection and data collection were performed independently by two researchers. PROSPERO registration: CRD42017062307.

Results Fifteen studies representing a total of 2986 patients were included. Inter-observer variation for CT-scan ranged from 11.2 to 69% (n = 678). Disagreement between ultrasound and CT-scan ranged between 6.6 and 17% (n = 221). Ten studies compared physical examination to CT-scan or ultrasound. Disagreement between physical examination and imaging ranged between 7.6 and 39% (n = 1602). Between 15 and 58% of IHs were solely detected by imaging (n = 483). Relative increase in IH prevalence for imaging compared to physical examination ranged from 0.92 to 2.4 (n = 1922).

Conclusions Ultrasound or CT-scan will result in substantial additional IH diagnosis. Lack of consensus regarding the definition of IH might contribute to the disagreement rates. Both the observer and diagnostic modality used could be additional factors explaining variability in IH prevalence and should be reported in IH research.

Keywords Incisional hernia · Diagnosis · Medical imaging · Hernia incidence

Introduction
Incisional hernia (IH) is the most frequent complication after open abdominal surgery. IH prevalence rates in published cohorts vary substantially: prevalence rates between 10 and 32% have been reported [1, 2]. Several factors explaining the variability in IH rate have been brought forward such as: age, obesity, abdominal aortic aneurysms, and previous abdominal surgery [1]. Most studies investigating the treatment or prevention of IH use IH prevalence as their primary endpoint. The diagnostic modality, observer, definition, and diagnostic protocol used for the diagnosis of IH are infrequently identified as factors associated with the IH prevalence rate. However, all four of these elements regularly differ within and between studies.
Many diagnostic modalities are used for the diagnosis of IH including physical examination, ultrasound, computed tomography scan (CT-scan), magnetic resonance imaging (MRI), and per-operative diagnosis. In IH research, the use of imaging modalities is considered important to achieve more reliable results. This is accentuated by the recommendation in the ‘European Hernia Society guidelines on the closure of abdominal walls’ to use ultrasound or CT-scan in the follow-up of prospective studies [3]. This approach deviates from everyday clinical practice, in which clinicians mainly focus on the diagnosis of symptomatic IHs that might require treatment [4].

In general, it is believed that the use of radiologic imaging will increase the detection rate of IH compared to physical examination alone. However, not all published cohorts show this trend [3–6].

The choice of diagnostic modality is often dictated by multiple factors such as cost, availability, safety, and especially in a research setting the detection rate, and reliability. However, the latter remains unclear, as the evidence concerning these factors is limited and sometimes contradictory [7, 8]. In IH research, the IH definition is not always uniform. The definition of IH as stated by Korienkov et al. [9]: ‘any abdominal wall gap with or without bulge in the area of a post-operative scar perceptible or palpable by clinical examination or imaging’, is acknowledged in the European Hernia Society (EHS) classification of primary and incisional abdominal wall hernias [9, 10]. Although IH is usually defined as an ‘abdominal wall gap or fascial defect’, some nuances with regard to this definition circulate as the term ‘abdominal wall weakness’ may also be used. Furthermore, bulging or a positive Valsalva maneuver may or may not be a diagnosing symptom [11, 12]. The place of imaging techniques within the diagnostic protocol often differs: some studies use a more clinical approach, reserving imaging techniques for cases with an inconclusive physical examination, whereas other studies only consider ‘radiologically confirmed’ diagnosis [2, 13, 14].

We hypothesize that the use of different diagnostic modalities, observers, definitions, and diagnostic protocols might influence the number of IHs identified. The objective of our systematic review is to evaluate the diagnostic accuracy of the different modalities used to identify IH after open abdominal surgery and after IH repair surgery. We provide a qualitative synthesis of the available data on the diagnostic accuracy of physical examination, CT-scan, and ultrasound for the identification of IH.

### Methods

The study protocol was registered in the PROSPERO database (International Prospective Register of Systematic Reviews, http://www.crd.york.ac.uk/prospero) prior to the start of the systematic review with the registration number CRD42017062307. All aspects of the PRISMA statement (Preferred Items for Reporting of Systematic Reviews and Meta-analyses) were followed [15].

### Search strategy

Embase, Medline ovid, Web of science, Cochrane, PubMed publisher, and Google scholar databases were searched on 28 March 2017. Full search details and syntax are presented in Appendix 1. The syntax construction and database search were performed in collaboration with a medical librarian specialized in conducting systematic reviews.

Studies reporting on IH diagnosis after primary laparotomy and after IH repair surgery were included. There was no limit in language or date of publication.

Studies were first evaluated for inclusion based on title and abstract by two independent researchers (LK and DS) and finally evaluated independently based on full text. Differences in article selection were discussed and articles were included or excluded after reaching agreement. Studies were included if they met the following criteria:

1. Inclusion of patients that underwent abdominal or IH repair surgery that were followed for the development of IH.
2. Studies assessing the performance of a diagnostic modality (physical examination, abdominal CT-scan, abdominal MRI scan, abdominal ultrasound, or surgery) used for the diagnosis of IH.

Studies assessing only laparoscopy patients, non-consecutive patient populations (e.g., patients with prior IH diagnosis), Spigelian, or occult hernias were excluded. Discrepancies in inclusion were resolved by discussion between reviewers and a senior author (JFL or FM).

### Data collection

Data collection was performed independently by two different researchers (LK and DS) using the standard forms covering study characteristics (study design, year, location, and level of evidence); patient baseline characteristics (type of intervention, number of patients, age, sex, open or laparoscopic surgery, duration of follow-up, and reason for surgery). Outcome characteristics concerning diagnostic performance comprise: definition of IH, inter-observer variation, CT-scan versus
ultrasound, CT-scan versus physical examination, ultrasound versus physical examination, diagnostic modalities versus per-operative diagnosis, and diagnostic performance in obese patients. Extracted data consisted of absolute data in four by four contingency tables, prevalence rates, kappa values, or intra-class correlation coefficients.

**Assessment of study quality**

The level of evidence of each paper was established according to the Oxford Centre for Evidence-based Medicine levels of evidence [16]. The possible risk of bias was assessed using the Cochrane Collaboration’s tool for assessing risk of bias [17]. Risk of bias was assessed separately for each outcome, since the quality of different outcomes in papers with a wide scope might differ.

**Results**

**Search and study characteristics**

The PRISMA flow diagram of the complete search strategy is shown in Fig. 1. The initial search resulted in 4855 articles (3010 after duplicates removal). After screening, 135 articles were selected for full-text reading. After full-text reading, 15 articles were selected for inclusion [2, 4–8, 11, 12, 14, 18–23]. Characteristics of included studies are summarized in Table 1.

**Study quality**

Risk of bias and applicability concerns of included studies per outcome are summarized in Fig. 2. Overall major concerns in patient selection, execution, and comparison of diagnostic tests and patient flow were present in 25–50% of the review sample (Fig. 3). Major applicability concerns were present in 10% of the review sample (Fig. 3). Specific methodological concerns are presented in Appendix 2.
A clear definition for IH was reported in seven of the included studies (Appendix 3) [2, 4, 7, 11, 12, 20, 22]. IH was defined as any ‘abdominal wall gap’ or ‘defect’ in the proximity of the post-operative scar, by five out of seven studies [2, 4, 7, 12, 22]. Two of these studies included ‘a protrusion of abdominal contents’ in the definition and incorporated the terms ‘weakness’ as well as ‘defect’ of the abdominal wall in their definition [12, 22].

**Definition of IH**

A clear definition for IH was reported in seven of the included studies (Appendix 3) [2, 4, 7, 11, 12, 20, 22]. IH was defined as any ‘abdominal wall gap’ or ‘defect’ in the proximity of the post-operative scar, by five out of seven studies [2, 4, 7, 12, 22]. Two of these studies included ‘a protrusion of abdominal contents’ in the definition and incorporated the terms ‘weakness’ as well as ‘defect’ of the abdominal wall in their definition [12, 22]. One study...
### Risk of Bias and Applicability Concerns Summary

| Study Description                  | Risk of Bias | Applicability Concerns |
|------------------------------------|--------------|------------------------|
|                                    | Patient Selection | Index Test | Reference Standard | Flow and Timing | Patient Selection | Index Test | Reference Standard |
| CT-scan vs Physical examination    |               |            |                      |                |               |            |                      |
| Baucom et al. 2014                | High         | High       | High                | Question        | High         | High       | High                |
| Caro-Tarrago et al. 2014          | High         | High       | High                | Question        | High         | High       | High                |
| Claes et al. 2014                 | High         | High       | High                | Question        | High         | High       | High                |
| Goodenough et al. 2015            | High         | High       | High                | Question        | High         | High       | High                |
| Gutiérrez de la Peña et al. 2001  | High         | High       | High                | Question        | High         | High       | High                |
| Holihan et al. 2016               | High         | High       | High                | Question        | High         | High       | High                |
| CT-scan vs ultrasound              |               |            |                      |                |               |            |                      |
| Beck et al. 2013                   | High         | High       | High                | Question        | High         | High       | High                |
| den Hartog et al. 2014             | High         | High       | High                | Question        | High         | High       | High                |
| Interobserver variation            |               |            |                      |                |               |            |                      |
| Baucom et al. 2016                 | High         | High       | High                | Question        | High         | High       | High                |
| Baucom et al. 2014                 | High         | High       | High                | Question        | High         | High       | High                |
| Claes et al. 2014                  | High         | High       | High                | Question        | High         | High       | High                |
| den Hartog et al. 2014             | High         | High       | High                | Question        | High         | High       | High                |
| Holihan et al. 2014                | High         | High       | High                | Question        | High         | High       | High                |
| Impact of obesity                  |               |            |                      |                |               |            |                      |
| Baucom et al. 2014                 | High         | High       | High                | Question        | High         | High       | High                |
| Bloemen et al. 2012                | High         | High       | High                | Question        | High         | High       | High                |
| Peroperative diagnosis             |               |            |                      |                |               |            |                      |
| Gutiérrez de la Peña et al. 2001  | High         | High       | High                | Question        | High         | High       | High                |
| Holihan et al. 2014                | High         | High       | High                | Question        | High         | High       | High                |
| Højer et al. 1997                  | High         | High       | High                | Question        | High         | High       | High                |
| Ultrasound vs physical examination |               |            |                      |                |               |            |                      |
| Baucom et al. 2016                 | High         | High       | High                | Question        | High         | High       | High                |
| Bloemen et al. 2012                | High         | High       | High                | Question        | High         | High       | High                |
| Deerenberg et al. 2015             | High         | High       | High                | Question        | High         | High       | High                |

**Legend:**
- **High:** High risk of bias
- **? Unclear:** Unclear risk of bias
- **Low:** Low risk of bias

**Fig. 2** Risk of bias and applicability concerns summary
defined IH as a ‘palpable protrusion’ under the laparotomy scar [11]. One study defined IH as ‘fascial defect’ in the proximity of the scar [20]. Three studies referred to a proposed universal definition [2, 4, 12]. One study that did not clearly define IH, reported that in case of disagreement between two or more observers, this was due to the lack of a clear definition among the observers in 35% of the patients (n = 42) [23].

**Inter-observer variation**

Inter-observer variation was reported in five of the included studies concerning a total of 698 patients [8, 12, 18, 20, 23]. Four out of five studies included in this comparison had one or more methodological concerns [12, 18, 19, 23]. Results obtained by these studies are summarized in Table 2. Reported disagreement between two observers ranged from 11.2 to 14.4%; corresponding kappa values ranged from 0.71 to 0.74 (n = 578) [8, 12, 18]. One study comparing the inter-observer variation in a group of six radiologists and three surgeons reported disagreement rates of 69 and 27%, respectively (kappa: 0.38 and 0.62; n = 100) [23]. One other study used a panel of five independent surgeons and reported an intra-class correlation coefficient of 0.85 (n = 20) [20]. The inter-observer variation of ultrasound was assessed in one study that used a panel of three independent surgeons, and an intra-class correlation coefficient of 0.79 (n = 17) was reported [7].

**Table 2** Inter-observer variation

| Study             | Risk of bias | Level of evidence | Radiologist B | Radiologist A |
|-------------------|--------------|-------------------|---------------|---------------|
| Den Hartog et al. [8] | +++          | 2B                | CT +          | CT −          | Total |
|                   | Agreement    | 87.50%            | CT +          | 21            | 1   | 22 |
|                   | Disagreement | 12.50%            | CT −          | 4             | 14  | 18 |
|                   | Kappa        | 0.74              | Total         | 25            | 15  | 40 |
| Baucom et al. [18] | ++           | 2B                | Surgeon       | Radiology report |
|                   | Agreement    | 85.60%            | CT +          | 78            | 21  | 99 |
|                   | Disagreement | 14.40%            | CT −          | 5             | 77  | 82 |
|                   | Kappa        | 0.71              | Total         | 83            | 98  | 181 |
| Claes et al. [12] | ++           | 2B                | Radiologist B | Radiologist A |
|                   | Agreement    | 88.80%            | CT +          | 84            | 21  | 105 |
|                   | Disagreement | 11.20%            | CT −          | 19            | 233 | 252 |
|                   | Kappa        | 0.73              | Total         | 103           | 254 | 357 |
| Holihan et al. [23] | ++          | N = 100           | Disagreement  | Kappa     |
|                   | Level of evidence | 2B              | 10 Observers  | 73          | 0.44 |
|                   | 10 Observers: 3 surgeons, 6 radiologist and radiology report | | 9 Observers | 71 | 0.44 |
|                   | Surgeons (n = 3) | | 27          | 0.62 |
|                   | Radiologists (n = 6) | | 69          | 0.38 |
| Baucom et al. [21] | +            | 3B                | Panel of 5 surgeons evaluated a random sample of 20 CT-scans. | Intra-class correlation coefficient: 0.85 |

![Fig. 3 Overall risk of bias and applicability concerns](image)
CT-scan versus ultrasound

The prevalence rate of IH after ultrasound and CT-scan was reported in two studies concerning a total of 221 patients [7, 8]. The study by Beck et al. [7] had methodological problems concerning patient selection and patient flow. Results obtained by these studies are summarized in Table 3. These two studies obtained contradictory results. Den Hartog et al. [8] reported a higher prevalence rate when using ultrasound, whereas Beck et al. [7] reported unchanged prevalence rates. Relative increase in prevalence rate when comparing CT-scan to ultrasound was 1.41 and 0.93. Disagreement between ultrasound and CT-scan was reported in 7/40 (17.5%) and 12/181 (6.6%) cases.

CT-scan versus physical examination

The prevalence rates of IH after CT-scan and physical examination were reported in six studies concerning a total of 1378 patients [5, 6, 11, 12, 14, 23]. Five out of six studies included in this comparison had one or more methodological concerns [5, 11, 12, 14, 23]. Results obtained by these studies are summarized in Table 4. Four studies reported higher prevalence rates and two studies reported lower prevalence rates when using CT-scan for the diagnosis of IH. The relative increase in prevalence rates when comparing ultrasound to physical examination ranged from 0.92 to 1.8 (n = 1013). Disagreement between diagnoses by ultrasound compared to physical examination was quantifiable in three studies. Disagreement between the two modalities was reported in 41/456 (9%), 44/338 (13%), and 15/38 (39%) of the cases. IH diagnosis was solely established with use of ultrasonography in 21/103 (20%), 41/87 (47%), and 15/26 (58%) of IH diagnosis [2, 4, 21].

Ultrasound versus physical examination

The prevalence rate of IH after ultrasound and physical examination was reported in four studies concerning a total of 1013 patients [2, 4, 7, 14, 21]. All studies included in this comparison had one or more methodological concerns [2, 4, 7, 14, 21]. Results obtained by these studies are summarized in Table 5. Three studies reported higher prevalence rates and one study reported a similar prevalence rate when using ultrasound for the diagnosis of IH. The relative increase in prevalence rates when comparing ultrasound to physical examination ranges from 1 to 2.4 (n = 1013). Disagreement between diagnoses by ultrasound compared to physical examination was quantifiable in three studies. Disagreement between the two modalities was reported in 41/456 (9%), 44/338 (13%), and 15/38 (39%) of the cases. IH diagnosis was solely established with use of ultrasonography in 21/103 (20%), 41/87 (47%), and 15/26 (58%) of IH diagnosis [2, 4, 21].

Per-operative diagnosis

The diagnosis obtained through physical examination or CT-scan was compared to the per-operative findings in three studies concerning 80 patients. Results obtained by these studies are summarized in Table 6 [6, 22, 23]. Only one of the studies included in this comparison was of good methodological quality. All reports on this outcome were flawed by small sample sizes. Gutiérrez de la Peña et al. [6] reported a true positive rate of 100% and a false positive rate of 98% (n = 50) for diagnosis with CT-scan. For the diagnosis with physical examination, a true positive rate of 75% and a false positive rate of 90% (n = 50) were reported [6].

Impact of obesity

The impact of obesity on the diagnosis of IH was reported in three studies concerning two different patient populations [4, 14, 19]. Baucom et al. [14] compared CT-scan as diagnostic modality to physical examination in obese and non-obese patients. The disagreement rate between the two modalities was 21% (n = 96) in obese patients.
compared to 13% in non-obese patients \((n = 85)\) [14]. Bloemen et al. [4] compared ultrasound as diagnostic modality to physical examination in patients with a body mass index (BMI) > 25 and in patients with a BMI < 25. The disagreement rate between the two modalities was 10% \((n = 228)\) in the BMI > 25 patients compared to 8% in BMI < 25 patients \((n = 228)\) [4]. One other study compared the mean surface area of incisional hernias detected with ultrasound in obese and non-obese patients and did not find a significant difference between the two [19].

### Discussion

In this systematic review on diagnostic modalities for IH diagnosis, great variance between modalities and between different studies was found. The diagnosis of IH remains challenging, as no objective gold standard is present. All included studies were of retrospective design, had multiple methodological concerns, or presented a small sample of patients (GRADE quality: low or very low). Therefore, the results of included studies should be interpreted with caution. Compared to per-operative diagnosis CT-scan seems to be reasonably accurate in one study presenting a small sample of patients [6]. However, considerable inter-observer variability has been reported [8, 12, 18, 20, 23]. Moreover, multiple studies report considerable discrepancy between CT-scan and physical examination and between CT-scan and ultrasonography results [2, 4–7, 11, 12, 14, 23]. No study compares ultrasound to the per-operative diagnosis. Two studies compare ultrasound to CT-scan and find contradictory results [7, 8]. Inter-observer variability for ultrasound and physical examination has not been assessed thoroughly; however, we may assume that inter-observer variability will be present due to the dynamic nature of these diagnostic modalities.

| Study                  | Risk of bias | Level of evidence | Prevalence PE | Prevalence CT | Relative increase | Total |
|------------------------|--------------|-------------------|---------------|---------------|------------------|-------|
| Gutierrez et al. [6]    | ++++         | 2B                | 18%           | 17%           | 0.92             | 5      |
| Baucom et al. [14]     | ++           | 2B                | 44%           | 55%           | 1.24             | 99    |
| Holihan et al. [23]    | ++           | 2B                | 30%           | 54%           | 1.80             | 101   |
| Goodenough et al. [5]  | ??           | 2B                | 18%           | 17%           | 0.92             | 360   |
| Caro-Tarrago et al. [11]| + +          | 2B                | 14%           | 20%           | 1.45             | 456   |
| Claes et al. [12]      | + +          | 2B                | 17%           | 30%           | 1.71             | 100   |

PE physical examination
### Table 5 Ultrasound versus physical examination

| Study Authors          | Risk of bias | Level of evidence | Prevalence PE | Prevalence US | Relative increase |
|------------------------|--------------|-------------------|---------------|---------------|------------------|
| Bloemen et al. [4]     | +++          | 2B                | 18.0%         | 18.2%         | 1.0              |
| Deerenberg et al. [2]  | ++           | 2B                | 13.6%         | 24.9%         | 1.8              |
| Baucom et al. [21]     | 3B           |                   | 28.9%         | 68.4%         | 2.4              |
| Baucom/Beck et al. [7, 14] | ++          |                   |               |               | n = 181          |

#### Table 6 Per-operative diagnosis

| Study Authors          | Risk of bias | Level of evidence | CT-scan + | CT-scan - | Surgery + | Surgery - | Total |
|------------------------|--------------|-------------------|-----------|-----------|-----------|-----------|-------|
| Gutiérrez de la Peña et al. [6] | +++          | 2B                | 8         | 1         | 1         | 1         | 9     |
| Hojer et al. [22]      | +++          | 3B                | 6         | 1         | 3         | 5         | 12    |
| Holihan et al. [23]    | +            | 3B                | 14        | 1         | 3         | 3         | 18    |

### Physical examination versus per-operative diagnosis

| Study Authors          | Risk of bias | Level of evidence | PE +      | PE -     | Surgery + | Surgery - | Total |
|------------------------|--------------|-------------------|-----------|----------|-----------|-----------|-------|
| Gutiérrez de la Peña et al. [6] | +++          | 2B                | 6         | 4        | 4         | 10        |       |
| Holihan et al. [23]    | +            | 3B                | 11        | 1        | 1         | 12        |       |
One prospective study of decent methodological quality provides a comparison between physical examination and the per-operative diagnosis in a small sample of 50 patients. Although the sample size was limited, this is the only report that provides some reliable insight in the sensitivity and specificity of physical examination, a sensitivity of 75%, and a specificity of 90% being reported [6]. Considerable discrepancies were reported between diagnoses by physical examination and ultrasound or CT-scan [2, 4–7, 11, 12, 14, 23]. Most studies report higher prevalence rates when using imaging modalities for the diagnosis of IH. However, not all studies show this trend [4, 6]. Relative increase in IH prevalence compared to physical examination ranged from 0.92 to 1.8 for CT-scan and 1 to 2.4 for ultrasound [2, 4–7, 11, 12, 14, 23]. Strikingly, studies that report similar prevalence rates for physical examination and ultrasound or CT-scan still show considerable disagreement between the two imaging modalities [4, 6]. The diagnostic performance of CT-scan is more thoroughly investigated compared to physical examination and ultrasound. CT-scan will likely provide the most sensitive and reproducible diagnosis of IH followed by ultrasound and physical examination. The definition of IH differed slightly in those studies that reported a definition. No study reported an IH definition specifically adapted for the diagnostic modality used. Disagreement between observers might in part be due to lack of consensus with regard to the IH definition [23].

It is important to stress that all the above-mentioned concerns relate to the research setting. For clinical studies, objective comparable measures should be used to report endpoints. The choice of diagnostic modality in a clinical setting might be relatively straightforward as most clinicians are mainly focused on identifying symptomatic incisional hernias that might require treatment. Therefore, in asymptomatic patients, a full diagnostic workup would often not be necessary. For a surgeon, detection rate is not the only argument to choose one modality over the other. In this case, costs, availability, patient safety, and patient comfort are important factors to take into account. It is understandable that a stepwise incremental approach is often chosen, in which physical examination will be the first modality used, followed by imaging in case of doubt.

In IH research, the diagnostic follow-up is challenging as no diagnostic gold standard exists and imaging will often be applied for non-IH related indications or in patients with an inconclusive physical examination, potentially causing for selection bias. The choice of diagnostic modality and the number of observers might influence the IH prevalence found. When different modalities and observers are unequally distributed over study cohorts, internal study validity could be compromised. This is especially of concern in studies of observational retrospective design, since many observers and different diagnostic modalities are present in every day clinical practice. Moreover, the aims of the clinician (identifying symptomatic IHs) often deviate from the aims of the researcher (identifying all IHs). Varying definitions for IH among observers are likely to cause a part of the observed disagreement [23].

Use of a universal definition such as the definition as proposed by Korenkov et al. [9]: ‘any abdominal wall gap with or without bulge in the area of a post-operative scar perceptible or palpable by clinical examination or imaging’, might be imperative. Based on current data, restricting the definition of IH to radiologically confirmed hernia’s only is not advisable, illustrated by the substantial inter-observer variation in CT-scan examinations and reports of false negative and false positive CT-scan diagnosis [6, 8, 18, 22, 23]. Although our knowledge with regard to inter-observer variation in IH diagnosis is mainly based on diagnosis by CT-scan, we may assume that these variations are of even more concern when applying ultrasound or physical examination, due to the more dynamic nature of these diagnostic modalities and the fact that in both modalities, subjectivity plays a larger role. The series presented by Holihan et al. [23] (CT-scan only) suggested that at least part of the observed inter-observer variation was due to subtle differences in the applied definition and methodology of operators. An IH definition specifically altered for the (radiologic) diagnostic modality of use, accompanied by a standardized systematic approach, might further improve the accuracy and consistency of IH diagnosis [7, 23]. For ultrasound examination, a systematic approach in which the midline area is examined first, followed by the abdominal areas next to the midline, and finally, the more lateral abdominal areas as suggested by Beck et al. [7] could be considered. This approach could be applied similarly for abdominal palpation. Since the diameter of the fascial defect and hernia sac significantly enlarges during a Valsalva maneuver, routine use of the Valsalva maneuver during physical examination, and radiologic evaluation of the post-operative scar might be of added diagnostic value [24].

The clinical relevance of IHs detected solely by radiologic imaging remains unclear. Only one study to date attempts to answer this question. Bloemen et al. [4] reported 26/103 of IH patients with discomfort. 3/26 of these IHs were detected by ultrasound alone, and 1/13 IHs that were treated surgically were detected by ultrasound alone. Based on current literature, the proportion of IHs solely detected by radiologic imaging that requires treatment or will progress through time remains unclear. Future research concerning the diagnosis of IHs should emphasize more on these factors.

**Limitations**

Our systematic review has some limitations. First, all included studies were of low quality: most were of
retrospective design, and some studies presented small samples. Therefore, the data should be interpreted with caution. We assume that between study, variation is present: follow-up, indication for abdominal surgery, BMI, and age differed between studies. In addition, some studies included a small proportion of laparoscopic patients [7, 12, 14, 18–20]. IH prevalence rates in patients operated laparoscopically differ from patients undergoing open abdominal surgery. Therefore, the proportion of patients operated laparoscopically will influence the total IH prevalence. Although these factors influence the comparability of reported IH prevalence, these factors might be of less concern when assessing the diagnostic accuracy. The majority of included studies had multiple methodological concerns. Risks for either reporting or selection bias was found frequently (Appendix 2). Most methodological concerns will mainly influence the overall prevalence rates; however, the diagnostic accuracy will be influenced by the prevalence rate to some degree. In addition, a number of studies did not compare the diagnostic modalities in a blinded fashion, potentially diluting the presented results and diminishing generalizability [2, 4, 5, 11, 12, 18].

Conclusion

Great variance between different diagnostic modalities and between different observers was found. Use of imaging modalities will usually cause for additional/increasing numbers of IH diagnosis and increase the IH prevalence compared to use of physical examination alone. When comparing different imaging modalities, CT-scan provides the most accurate diagnosis. Lack of consensus with regard to the IH definition among observers might in part explain the inter-observer variation. The observer, diagnostic modality, and diagnostic approach could be additional factors explaining variability in IH prevalence and should, therefore, be reported with detail in IH research. To achieve internally valid study results, proper distribution of different observers and diagnostic modalities across study cohorts is imperative.

Acknowledgements We would like to thank Wichor Bramer for his assistance on the search strategy and syntax.

Compliance with ethical standards

Conflict of interest LFK, DS, GJK, and JFL declare that they have no conflict of interest. FM declares conflict of interest not related to the submitted work, grants, and personal fees from Medtronic and Dynamesh.

Ethical approval This study did not need approval from the local ethical committee.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

Informed consent For this type of article informed consent is not required.

Open Access This article is distributed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (http://creativecommons.org/licenses/by-nc/4.0/), which permits any noncommercial use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Appendix 1: Literature search syntax

Embase.com

('incisional hernia'/exp OR 'abdominal wall hernia'/mj OR ((incision* OR scar* OR cicatri*) NEAR/3 (herni*)) OR post-operat*-herni* OR post-operat*-herni*):ab,ti OR ((abdom* OR ventral*) NEAR/3 (herni*)):ti AND ('sensitivity and specificity'/exp OR 'diagnostic value'/exp OR 'interrater reliability'/exp OR 'reproducibility'/de OR 'observer variation'/exp OR 'observer bias'/exp OR 'diagnostic error'/exp OR 'diagnostic accuracy'/de OR 'diagnostic test accuracy study'/exp OR 'differential diagnosis'/exp OR 'predictive value'/de OR 'kappa statistics'/de OR (sensitiv* OR specific* OR ((diagnos* OR imaging OR ct OR tomograph* OR resonance OR mri OR predicti*):ab,ti OR ((('diagnosis'/de OR 'computer-assisted diagnosis'/exp OR 'diagnosis':lnk OR 'imaging and display'/exp OR 'computer-assisted tomography'/exp OR 'nuclear magnetic resonance imaging'/exp OR radiodiagnos*de OR 'diagnostic imaging'/exp OR tomography/exp OR 'nuclear magnetic resonance'/exp OR 'physical examination'/exp OR 'ultrasound'/de OR 'echography'/exp OR 'Valsalva maneuver'/de OR 'patient-reported outcome'/exp OR (diagnos* OR radiodiagnos* OR misdiagnos* OR imaging OR (compute* NEAR/3 tomogra*) OR ((ct OR cat OR mr OR nmr) NEXT/3 (scan* OR imag*)):OR mri OR (magnet* NEAR/3 resonan*):OR (physical* NEAR/3 examinat*) OR ultraso* OR sonogra* OR echogra* OR patient-report* OR palpat* OR Valsalva):ab,ti) AND ('intermethod comparison'/exp OR 'comparative study'/de OR 'instrument validation'/de OR 'validation process'/de OR 'validation study'/de OR

Springer
(compare* OR comparative* OR comparison* OR comparing* OR validat* OR evaluat*):ab,ti)).

**Medline Ovid**

("Incisional Hernia"/ OR "Hernia, Ventral"/ OR ((incision* OR scar* OR cicatri*) ADJ3 (herni*)):ab,ti,kf. OR (abdom* OR ventral*) ADJ3 (herni*)):ti.) AND ("Sensitivity and Specificity"/ OR "Reproducibility of Results"/ OR "observer variation"/ OR exp "diagnostic errors"/ OR "Diagnosis, Differential"/ OR "kappa statistics"/ OR (sensitiv* OR specific* OR ((diagnos* OR imaging OR ct OR tomograph* OR resonance OR mri OR predicti*) ADJ6 (value* OR useful* OR challenge* OR pitfall* OR contribution* OR effect* OR efficac* OR error* OR accura* OR different*)): OR (false ADJ (negative* OR positive*)): OR ppv OR npv OR reliab* OR reprodu* OR interrat* OR observer* OR inter-observer* OR intraobserver* OR (kappa ADJ (value OR test OR statistic*)):ab,ti,kf. OR ("diagnosis"/ OR exp "Diagnosis, Computer-Assisted"/ OR "diagnosis": xs. OR exp "Magnetic Resonance Imaging"/ OR exp "diagnostic imaging": OR exp tomography/ OR "Magnetic Resonance Spectroscopy"/ OR exp "physical examination": OR "Ultrasonics": OR exp "Ultrasonography": OR "Valsalva Maneuver": OR "Patient-Reported Outcome Measures": OR ((diagnos* OR radiodiagnos* OR misdiagnos* OR imaging OR (compute* ADJ3 tomogra*) OR (ct OR cat OR mr OR nmr) ADJ3 (scan* OR imag*)): OR mri OR (magnet* ADJ3 resonan*): OR (physical* ADJ3 examinat*): OR ultraso* OR sonogra* OR echogra* OR patient-report* OR palpat* OR Valsalva):ab,ti,kf.) AND ("Comparative Study": OR "Validation Studies": OR "evaluation studies": OR (compare* OR comparative* OR comparing* OR validat* OR evaluat*):ab,ti,kf.)).

**Web of science**

TS=((((incision* OR scar* OR cicatri*) NEAR/2 (herni*)) OR post-operat*-herni* OR post-operat*-herni*)): AND ("sensitiv* OR specific* OR ((diagnos* OR imaging OR ct OR tomograph* OR resonance OR mri OR predicti*) NEAR/5 (value* OR useful* OR challenge* OR pitfall* OR contribution* OR effect* OR efficac* OR error* OR accura* OR different*)): OR (false NEAR/1 (negative* OR positive*)): OR ppv OR npv OR reliab* OR reprodu* OR interrat* OR observer* OR inter-observer* OR intraobserver* OR (kappa NEAR/1 (value OR test OR statistic*)): OR (((diagnos* OR radiodiagnos* OR misdiagnos* OR imaging OR (compute* NEAR/2 tomogra*): OR (ct OR cat OR mr OR nmr) NEAR/2 (scan* OR imag*)): OR mri OR (magnet* NEAR/2 resonan*): OR (physical* NEAR/2 examinat*): OR ultraso* OR sonogra* OR echogra* OR patient-report* OR palpat* OR Valsalva)) AND (compare* OR comparative* OR comparing* OR validat* OR evaluat*):)).

**Google scholar**

"incisional|scar|cicatrical hernia" diagnosis|radiodiagnosis|slimming|tomograph|mri" physical examination"ultraso nographylechography validation|sensitivity|specificity|diagnostic value|error|accuracy".

**Appendix 2: Methodological concerns**

| Inter-observer variation | Den Hartog et al. [8] | No major methodological concerns |
|--------------------------|-----------------------|----------------------------------|
| Baucom et al. [18]       | Surgeon was asked to specifically diagnose incisional hernia, radiologists were not (reporting bias) |
| Claes et al. [12]        | Inclusion of small proportion of laparoscopic patients (applicability) |
| Holihan et al. [23]      | Two radiologists report higher prevalence rates potentially diluting the results (reporting bias) |
| Study                          | Description                                                                                     |
|-------------------------------|-------------------------------------------------------------------------------------------------|
| Baucom et al. [20]            | Small sample out of larger cohort (selection bias)                                             |
| Den Hartog et al. [8]         | No major methodological concerns                                                                |
| Beck et al. [7]               | Patients were only included if a CT-scan was available (selection bias); interval between CT-scan and ultrasound up to 6 months (reporting bias) |
| Gutiérrez de la Peña et al. [6] | No major methodological concerns                                                              |
| Holihan et al. [23]          | Only patients with an available CT-scan were included, the surgically evacuated patients consist of a non-random sample; decision to operate was made based on CT-scan (selection bias) |
| Caro-Tarrago et al. [11]     | Interpretation of CT-scan not blinded to results of physical examination (reporting bias)      |
| Claes et al. [12]            | Patient samples differed per modality (selection bias); interval between CT-scan and physical examination unclear (reporting bias); unclear whether comparison was made blinded (reporting bias) |
| Bloemen et al. [4]           | Ultrasound was not performed blinded to the results of physical examination (reporting bias)   |
| Deerenberg et al. [2]        | Unclear whether comparison was made blinded (reporting bias)                                  |
| Baucom et al. [21]           | Included patients were likely to have an incisional hernia, high losses to follow-up related to outcome (selection bias); comparison between modalities was not blinded (reporting bias) |
| Baucom/Beck et al. [7, 14]   | Patients were only included if a CT-scan was available (selection bias)                        |
| Højer et al. [22]            | Patients selected for this study had an inconclusive physical examination for incisional hernia (selection bias) |
| Gutiérrez de la Peña et al. [6] | No major methodological concerns                                                              |
| Holihan et al. [23]          | Only patients with an available CT-scan were included, the surgically evacuated patients consist of a non-random sample, decision to operate was made based on CT-scan (selection bias); physical examination results were obtained through patient records (reporting bias) |

**Appendix 3: Incisional hernia definition**

| Study                          | Definition of incisional hernia                                                                 |
|-------------------------------|-------------------------------------------------------------------------------------------------|
| Baucom et al. [21]            | ‘…any fascial defect within 7 cm of an incision made at the time of the cancer operation’    |
| Beck et al. [7]               | ‘Full-thickness defect in the abdominal wall fascia or lateral muscular […] in the region of a previous incision’ |
| Bloemen et al. [4]            | ‘Any abdominal wall gap with or without a bulge in the area of a post-operative scar, palpable or perceptible by clinical examination or imaging’ |
| Caro-Tarango et al. [11]     | ‘…a palpable hernial protrusion under the laparotomy scar when Valsalva manoeuvres were carried out in the supine decubitus position and/or in the bipedestacon posture’ |
| Claes et al. [12]            | ‘An abnormal protrusion of the contents of the abdominal cavity or of pre-peritoneal fat through a defect or weakness in the abdominal wall at the site of the surgical scar’ |
| Deerenberg et al. [2]        | ‘any abdominal wall gap with or without bulge in the area of a post-operative scar perceptible or palpable by clinical examination or imaging’ |
| Højer et al. [22]            | ‘…a peritoneal sac that protrudes through a weakness or defect in the muscular and fascial layers of the abdomen’ |
References

1. Bosanquet DC, Ansell J, Abdelrahman T, Cornish J, Harries R, Stimpson A, Davies L, Glaseby JCD, Frewer KA, Frewer NC (2015) Systematic review and meta-regression of factors affecting midline incisional hernia rates: analysis of 14,618 patients. PLoS One 10(9):e0138745

2. Deerenberg EB, Harlaar JJ, Steyerberg EW, Lont HE (2015) Small bites versus large bites for closure of abdominal midline incisions (STITCH): a double-blind, multicentre, randomised controlled trial. Lancet 386(10000):1254–1260

3. Muysoms FE, Antoniou SA, Bury K, Campanelli G, Conze J, Cucurullo d, De Beaux AC, Deerenberg EB, Harlaar JJ, Steyerberg EW, Lont HE (2015) Small bites versus large bites for closure of abdominal midline incisions (STITCH): a double-blind, multicentre, randomised controlled trial. Lancet 386(10000):1254–1260

4. Bloemen A, Van Dooren P, Huizinga BF, Hoofwijk AGM (2012) Comparison of ultrasonography and physical examination in the diagnosis of incisional hernia in a prospective study. Hernia 16(1):53–57. https://doi.org/10.1007/s10029-011-0865-2

5. Goodenough CJ, Ko TC, Kao LS, Nguyen MT, Holihan JL, Alawadi Z, Nguyen DH, Flores JR, Arita NT, Roth JS, Liang MK (2015) Development and validation of a risk stratification score for ventral incisional hernia after abdominal surgery: Hernia expectation rates in intra-abdominal surgery (The HERNIA project). J Am Coll Surg 220(4):405–413. https://doi.org/10.1016/j.jamcollsurg.2014.12.027

6. de la Peña CG, Romero JV, Garcia JAD (2001) The value of CT hernia recurrence after prophylactic repair of ventral incisional hernias. Eur Radiol 11(7):1161–1164. https://doi.org/10.1007/s003300010743

7. Beck WC, Holzman MD, Sharp KW, Nealon WH, Dupont WD, Poulose BK (2013) Comparative effectiveness of dynamic abdominal Sonography for hernia vs computed tomography in the diagnosis of incisional hernia. J Am Coll Surg 216(3):447–453. https://doi.org/10.1016/j.jamcollsurg.2012.11.012

8. den Hartog D, Dur AHM, Kamphuis AGA, Tuinebeijer WE, Kreis RW (2009) Comparison of ultrasonography with computed tomography in the diagnosis of incisional hernias. Hernia 13(1):45–48. https://doi.org/10.1007/s10029-008-0420-y

9. Korenkov M, Paul A, Sauerland S, Neugebauer E, Arndt M, Chevrel JP, Corcione F, Fingert A, Flamet JB, Kux M, Matzinger A, Myrvold HE, Rath AM, Simmermacher RK (2001) Classification and surgical treatment of incisional hernia. Results of an experts’ meeting. Langenbecks Arch Surg 386(1):65–73

10. Muysoms FE, Miserez M, Berrevoet F, Campanelli G, Champault GG, Chelala E, Dietz UA, Eker HH, El Nakadi I, Hauters P (2009) Classification of primary and abdominal incisional wall hernias. Hernia 13(4):407–414

11. Caro-Tarrago A, Casas CO, Salido AJ, Guileda ER, Fernandez FM, Guillen VV (2014) Prevention of incisional hernia in midline laparotomy with an onlay mesh: a randomized clinical trial. World J Surg 38(9):2223–2230. https://doi.org/10.1007/s00268-014-2510-6

12. Claes K, Beckers R, Heindryckx E, Kyle-Leinhas I, Pletinckx P, Claey D, Muysoms F (2014) Retrospective observational study on the incidence of incisional hernias after colorectal carcinoma resection with follow-up CT scan. 18:787–802. https://doi.org/10.1007/s10029-014-1214-z

13. Bloemen A, Van Dooren P, Huizinga BF, Hoofwijk AGM (2011) Randomized clinical trial comparing polypropylene or polydioxanone for midline abdominal wall closure. Br J Surg 98(5):633–639. https://doi.org/10.1002/bjs.7398

14. Baucom RB, Beck WC, Holzman MD, Sharp KW, Nealon WH, Poulose BK (2014) Prospective evaluation of surgeon physical examination for detection of incisional hernias. J Am Coll Surg 218(3):363–366. https://doi.org/10.1016/j.jamcollsurg.2013.12.007

15. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma G (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med 6(7):e1000097

16. Interval C (2010) Oxford Centre for Evidence-based Medicine Levels of Evidence (May 2001)

17. Higgins JPT, Green S (2011) Cochrane handbook for systematic reviews of interventions, vol 4. Wiley, New York

18. Baucom RB, Beck WC, Holzman MD, Sharp KW, Nealon WH, Poulose BK (2014) The importance of surgeon-reviewed computed tomography for incisional hernia detection: a prospective study. Am Surg 80(7):720–722

19. Baucom RB, Beck WC, Phillips SE, Holzman MD, Sharp KW, Nealon WH, Poulose BK (2014) Comparative evaluation of dynamic abdominal sonography for hernia and computed tomography for characterization of incisional hernia. JAMA Surg 149(6):591–596. https://doi.org/10.1001/jamasurg.2014.36

20. Baucom RB, Ousley J, Beveridge GB, Phillips SE, Pierce RA, Holzman MD, Sharp KW, Nealon WH, Poulose BK (2016) Cancer survivorship: defining the incidence of incisional hernia after resection for intra-abdominal malignancy. Ann Surg Oncol 23:764–771. https://doi.org/10.1245/s10434-016-5546-z

21. Baucom RB, Ousley J, Feurer ID, Beveridge GB, Pierce RA, Holzman MD, Sharp KW, Poulose BK (2016) Patient reported outcomes after incisional hernia repair—establishing the ventral hernia recurrence inventory. Am J Surg 212(1):81–88. https://doi.org/10.1016/j.amjsurg.2015.06.007

22. Hejser AM, Rygaard H, Jess P (1997) CT in the diagnosis of abdominal wall hernias: a preliminary study. Eur Radiol 7(9):1416–1418

23. Holihan JL, Karanjavala B, Ko A, Askensy EP, Matta EJ, Gharbouali L, Hasapess J, Tammisetti VS, Thupili CR, Alawadi ZM, Bondre I, Flores-Gonzalez JR, Kao LS, Liang MK (2016) Use of computed tomography in diagnosing ventral hernia recurrence: a blinded, prospective, multispecialty evaluation. JAMA Surg 151(1):7–13. https://doi.org/10.1001/jamasurg.2015.2580

24. Jaffe TA, O’Connell MJ, Harris JP, Paulson EK, DeLong DM (2005) MDCT of abdominal wall hernias: is there a role for Val-salva’s maneuver? Am J Roentgenol 184(3):847–851